UNIVERSITY OF ALBERTA

REGULATION OF CEPHAMYCIN BIOSYNTHESIS IN *STREPTOMYCES CLAVULIGERUS*

by

DYLAN CHRISTOPHER ALEXANDER ©

A thesis submitted to the faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY

in

MICROBIOLOGY AND BIOTECHNOLOGY

DEPARTMENT OF BIOLOGICAL SCIENCES

EDMONTON, ALBERTA
FALL, 1998

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-34723-0

Canada

UNIVERSITY OF ALBERTA

LIBRARY RELEASE FORM

Name of Author:	DYLAN CHRISTOPHER ALEXANDER
Title of Thesis:	REGULATION OF CEPHAMYCIN BIOSYNTHESIS IN *STREPTOMYCES CLAVULIGERUS*
Degree:	DOCTOR OF PHILOSOPHY
Year this Degree Granted:	1998

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Dylan Alexander

DYLAN C. ALEXANDER
18523-82 AVE
Edmonton, Alberta
CANADA
T5T 1G2

Date: June 22, 1998

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled REGULATION OF CEPHAMYCIN BIOSYNTHESIS IN *STREPTOMYCES CLAVULIGERUS* submitted by DYLAN CHRISTOPHER ALEXANDER in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in MICROBIOLOGY AND BIOTECHNOLOGY.

SE Jensen
Dr. S. E. Jensen-Supervisor

Laura Frost
Dr. L. S. Frost

Brenda Leskiw
Dr. B. K. Leskiw

J N Mark Glover
Dr. M. Glover

CR Hutchinson
Dr. C. R. Hutchinson

Date: June 22 1998

ABSTRACT

In general, antibiotic biosynthesis in *Streptomyces* is transcriptionally regulated via pathway-specific activators. To determine whether this type of activator may be involved in the regulation of cephamycin biosynthesis in *Streptomyces clavuligerus*, sequencing of the cephamycin biosynthetic cluster was completed. The *ccaR* gene, which displays homology to many *Streptomyces* pathway-specific transcriptional activators, was further characterized to confirm its role in the regulation of cephamycin biosynthesis. As well, the genetic organization of the first three biosynthetic genes was characterized.

Sequence analysis of previously unsequenced gaps in the cluster revealed the presence of eight new genes. These include genes encoding an apparent penicillin binding protein and a transport protein, as well as three proteins which catalyze late reactions in the biosynthetic pathway. Also found were two genes with no obvious function in cephamycin biosynthesis that may define the end of the cluster. Finally, the *pcd* gene which displays significant homology to semialdehyde dehydrogenase genes may represent the long-sought-after dehydrogenase catalyzing the second of two steps converting lysine to α-aminoadipate.

Mutation of the *ccaR* gene determined that CcaR was essential for the biosynthesis of cephamycin, clavulanic acid and clavams. Complementation of a deletion mutant lacking the *ccaR*, *orf11* and *blp* genes showed that only *ccaR* was essential for the biosynthesis of all these β-lactam metabolites. The lack of cephamycin production in *ccaR* mutants was directly attributable to the absence of biosynthetic enzymes responsible for the early and middle steps of the cephamycin biosynthetic pathway. Complementation of the *ccaR* deletion mutant returned these biosynthetic enzymes and restored cephamycin production.

The genes for LAT, ACVS and IPNS proteins are encoded on the *lat/pcbAB/pcbC*

polycistronic transcript that initiates from the *lat* promoter. Interruption of the *lat* gene within the polycistronic transcript with an antibiotic resistance marker equally decreases ACVS and IPNS protein production. Termination of the polycistronic transcript with a transcription terminator eliminates production of the ACVS and IPNS proteins. The *pcbC* gene has its own promoter which creates a *pcbC* monocistronic transcript but the absence of the IPNS protein suggests that this transcript can not be translated.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to Dr. Jensen for her support and guidance during my years in her lab. Her patience has not gone unnoticed. I would also like to thank Dr. Leskiw, Dr. Roy and Dr. Frost for their experience, guidance and respect throughout my years.

Others whom I would like to thank have taught me techniques and/or provided fruitful advice on how to be successful in the lab. These people include Tim van Biesen, Astrid Petrich, Annie Wong, Karen Anthony, Barry Barton and Leigh Ann Giebelhaus and I'm sure many others whom I've forgotten but were nevertheless appreciated. Leah Williams and Linda Lee were summer students who provided many hours of dedication and hard work towards the completion of this thesis and for that I would like to thank them also.

To my family; my parents, grandmother, aunts, uncles and cousins, and to my friends I thank you for your support and time spent outside of the lab. It too has contributed to this thesis.

I would also like to acknowledge the support through graduate studentships from the Alberta Heritage Foundation for Medical Research and the Natural Sciences and Engineering Research Council.

TABLE OF CONTENTS

Page

I. Introduction 1
  I.1 β-lactam antibiotics 1
  I.2 Clavulanic acid/clavam biosynthetic pathways and clusters 4
  I.3 Penicillin/cephalosporin/cephamycin biosynthetic pathway 9
  I.4 Cephamycin biosynthetic cluster 14
  I.5 Organization and regulation of the *lat*, *pcbAB* and *pcbC* genes 18
  I.6 Regulation of antibiotic production in *Streptomyces* 29
  I.7 Regulatory genes within the *S. clavuligerus* cephamycin cluster 38
II. Materials and Methods 40
  II.1 Materials 40
  II.2 *Streptomyces* procedures 41
    II.2.1 Storage of strains 41
    II.2.2 Introduction of DNA into *Streptomyces* 44
      II.2.2.1 Preparation of protoplasts from *S. clavuligerus* strains 44
      II.2.2.2 Transformation of *S. clavuligerus* 44
    II.2.3 Analysis of DNA from *Streptomyces* 46
      II.2.3.1 Plasmid DNA isolation 46
      II.2.3.2 Genomic DNA isolation 46
      II.2.3.3 Southern blotting 47
      II.2.3.3.1 Hybridization of Southern blots 47
    II.2.4 *S. clavuligerus* fermentation culture conditions and media 48
      II.2.4.1 Harvesting and preparation of cell-extracts 49
      II.2.4.2 Analysis of cell-extracts for biosynthetic enzyme activities 49
        II.2.4.2.1 δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase assay 49
        II.2.4.2.2 Isopenicillin N synthase bioassay 50
        II.2.4.2.3 Isopenicillin N epimerase bioassay 51
        II.2.4.2.4 Desacetoxycephalosporin C synthase bioassay 51
      II.2.4.3 Western blot analysis of cell-extracts 52
        II.2.4.3.1 Quantitation of protein content in cell-extracts 52
        II.2.4.3.2 SDS-PAGE for ACVS protein detection 52
        II.2.4.3.3 SDS-PAGE and transfer to PVDF membranes 53
        II.2.4.3.4 Developing western blots 53
      II.2.4.4 Quantitation of antibiotic production and growth 54

II.2.4.4.1 Bioassay for penicillin/cephamycin 54
II.2.4.4.2 HPLC analysis for clavulanic acid and other clavams 54
II.2.4.4.3 Growth determination 55
II.2.4.5 Harvesting cultures for *xylE* promoter probe analysis 55
II.2.4.5.1 Catechol 2,3 dioxygenase assays 55
II.2.5 Generation of gene replacement mutants by homologous recombination 56
II.3 *E. coli* procedures 57
II.3.1 Storage of strains 57
II.3.2 Culture conditions and media 57
II.3.3 Introduction of DNA into *E. coli* 58
II.3.3.1 Preparation of electrocompetent *E. coli* cells 58
II.3.3.2 Electroporation of DNA into electrocompetent *E. coli* 58
II.3.3.3 Preparation of $MnCl_2$ competent *E. coli* cells 58
II.3.3.4 Transformation of DNA into $MnCl_2$ competent *E. coli* 59
II.3.4 Analysis of DNA from *E. coli* 59
II.3.4.1 Plasmid DNA isolation 59
II.3.4.2 Southern blotting and hybridization 59
II.3.4.3 Colony hybridization 60
II.3.4.4 DNA sequence analysis 60
II.3.5 Production of MBP-fusion proteins with LAT or CcaR in *E. coli* to create polyclonal antibodies in rabbits 61
II.3.5.1 Production of MBP-fusion proteins 61
II.3.5.2 Purification of MBP-fusion proteins 61
II.3.5.3 Preparation of antigen for injection into rabbits 78
II.3.5.4 Collection of polyclonal antibody containing sera 78
II.3.5.5 Purification of antibodies from western blots 78
II.3.5.5.1 Production and purification of GST-fusion proteins 78
II.3.5.5.2 Immunoaffinity purification of antibodies 79
II.4 Recombinant DNA Techniques 79
II.4.1 Organic solvent extraction of DNA and ethanol precipitation 79
II.4.2 Gel purification of DNA fragments 81
II.4.3 Dephosphorylation of DNA ends 81
II.4.4 Generation of blunt-ends with Klenow or T4 DNA polymerase 82
II.4.5 Random primer labeling of hybridization probes 82
II.4.6 Ligation of DNA Molecules 82
II.4.7 Creation of exoIII deletion clones 83

II.4.7.1 Generation of exoIII deletions for sequencing 83
II.4.7.1.1 Screening deletion clones for sequencing 85
II.4.7.2 ExoIII deletions in the 3' end of the *ccaR* gene 85
II.4.8 PCR procedures 85
II.4.8.1 PCR of the *ccaR* coding region for expression in *E. coli* 85
II.4.8.2 PCR of the *lat* coding region for expression in *E. coli* 86
II.4.8.3 PCR of the pSET152 MCS for clone confirmation 87
II.4.8.4 PCR of the pDA2000 MCS for clone confirmation 87
III. Results I. Sequencing and analysis of the cephamycin cluster 89
III.1 Cosmid mapping and subcloning of cluster fragments 89
III.2 Sequencing of the cephamycin cluster 93
III.3 Organization of sequence data from the cephamycin cluster 95
III.4 Analysis of the cephamycin cluster sequence information 112
III.5 Characterization of the cephamycin cluster gene products 116
III.5.1 The *rhsA* gene product 119
III.5.2 The *sclU* gene product 120
III.5.3 The *bla* gene product 120
III.5.4 The *pbpA* gene product 121
III.5.5 The *cmcT* gene product 122
III.5.6 The *pcd* gene product 122
III.5.6.1 Alignment of the PCD protein active site with homologs 122
III.5.7 The *cefE* gene product 125
III.5.8 The *cefD* gene product 125
III.5.9 The *cmcI* gene product 128
III.5.10 The *cmcJ* gene product 128
III.5.11 The *cefF* gene product 128
III.5.12 The *cmcH* gene product 128
III.5.13 The *ccaR* gene product 129
III.5.14 The *orf11* gene product 129
III.5.15 The *blp* gene product 129
III.5.15.1 Alignment of Blp protein with BLIP 129
III.5.16 The *lat* gene product 130
III.5.17 The *pcbAB* gene product 130
III.5.18 The *pcbC* gene product 130
III.5.19 The *pcbR* gene product 130
III.6 Characterization of the CcaR protein 133

III.6.1 Detection of proteins homologous to CcaR 133
III.6.2 Alignment of CcaR with homologous proteins 133
III.6.3 Estimation of DclX/CcaR start codon 137
III.6.4 Detection of structural motifs in the CcaR sequence 137
III.7 Comparison of the *N. lactamdurans* and *S. clavuligerus* cephamycin biosynthetic clusters 140
III.8 Heterologous probing of actinomycetes 143
III.8.1 Hybridization with *S. clavuligerus* DNA fragments 144
IV. Results. Regulation of the *lat*, *pcbAB* and *pcbC* genes 149
IV.1 Construction of promoter probe shuttle vectors 150
IV.1.1 Reconstructing the reporter gene cassette from pIJ4083 151
IV.1.2 Construction of a mid-copy number promoter probe vector 151
IV.1.3 Construction of a single copy promoter probe vector 151
IV.2 Cloning promoter fragments into pDA2000 or pDA2500 156
IV.2.1 Verification of the pDA2000-based plasmid integration by genomic PCR 161
IV.3 Promoter probe analysis of the *lat* gene 161
IV.3.1 Construction of pDA2000-based promoter probe plasmids with *lat* gene fragments 161
IV.3.2 Promoter probe analysis in wildtype *S. clavuligerus* 166
IV.4 Alteration of the *lat* gene by gene replacement 168
IV.4.1 Partial *lat* deletion and replacement with the *tsr* marker (Δ*lat*::*tsr* mutants) 169
IV.4.2 Partial *lat* deletion and replacement with the *tsr* marker and FKMT terminator (Δ*lat*::*tsr*/term mutants) 174
IV.4.2.1 Construction of the FKMT-containing promoter probe plasmids 177
IV.4.2.2 Promoter probe analysis of the effectiveness of the FKMT terminator 177
IV.4.3 Partial *lat* deletion after removal of the *apr* marker from *lat*::*apr* (Δ*lat* mutants) 178
IV.4.4 Additional Southern blot analyses of *lat* mutant strains 183
IV.5 Characterization of *lat* mutants created by gene replacement 183
IV.5.1 Characterization of cephamycin C production by the *lat* mutants 184
IV.5.2 Western blot analysis of the *lat* mutants 184
IV.5.2.1 Western blot analysis of the Δ*lat*::*tsr* mutants 186
IV.5.2.2 Western blot analysis of the Δ*lat*::*tsr*/term mutants 186
IV.5.2.3 Western blot analysis of the Δ*lat* mutants 189
IV.5.3 Comparison of cephamycin biosynthetic enzyme activities in the various *lat* mutants 189

IV.5.3.1 Biochemical characterization of the *lat* mutants 189
IV.5.3.2 Western blot analysis of the *lat* mutants 191
IV.5.3.3 Analysis of the *lat* mutants for production of ACVS by SDS-PAGE 191
IV.6 Complementation of *lat* mutants 194
IV.6.1 Verification of the pSET152-based plasmid integration by genomic PCR 194
IV.6.2 Characterization of cephamycin C production by the complemented *lat* mutants 199
IV.6.3 Western blotting analysis of the complemented *lat* mutants 201
IV.6.4 Analysis of the complemented *lat* mutants for production of ACVS by SDS-PAGE 201
IV.7 Promoter probe analysis of *lat* mutants 206
IV.7.1 Analysis of *lat* promoter activity in *lat* mutants 206
V. Results. Regulation of β-lactam biosynthesis by the CcaR protein 208
V.1 Mutation of the *ccaR* gene or neighboring genes by gene replacement 208
V.1.1 Construction of the *apr* cassette 209
V.1.2 Disruption of the *cob* genes with the *apr* cassette 210
V.1.2.1 *ccaR* disruption with the *apr* cassette (*ccaR*::*apr*) 210
V.1.2.2 *orf11* disruption with the *apr* cassette (*orf11*::*apr*) 215
V.1.2.3 *blp* disruption with the *apr* cassette (*blp*::*apr*) 218
V.1.3 Deletion of *cob* genes and replacement with the *tsr* marker 218
V.1.3.1 *ccaR* deletion and replacement with the *tsr* marker (Δ*ccaR*::*tsr*) 219
V.1.3.2 *cob* deletion and replacement with the *tsr* marker (Δ*cob*::*tsr*) 219
V.1.4 Additional Southern blot analyses of mutant strains 226
V.2 Characterization of *ccaR* mutants 226
V.2.1 Characterization of the cephamycin C production by the *ccaR* mutants 227
V.2.2 Western blot analysis of the *ccaR* mutants 227
V.2.2.1 Western blot analysis of the *ccaR*::*apr* mutants 227
V.2.2.2 Western blot analysis of the Δ*ccaR*::*tsr* mutants 227
V.2.2.3 Western blot analysis of the Δ*cob*::*tsr* mutants 230
V.3 Characterization of the NTG mutants 230
V.3.1 Production of cephamycin C by the NTG mutants 231
V.3.2 Western blot analysis of the NTG mutants 231
V.4 Comparison of cephamycin biosynthetic activities in the various *ccaR* mutants 231
V.4.1 Biochemical characterization of the *ccaR* mutants 231
V.4.2 Western blot analysis of the *ccaR* mutants 234
V.4.3 Analysis of the *ccaR* mutants for production of ACVS by SDS-PAGE 234

V.5. Characterization of the *orf11*::*apr* mutants 234
  V.5.1 Production of cephamycin C by *orf11*::*apr* mutants 237
V.6 Production of other β-lactam antibiotics by mutants strains 237
  V.6.1 Characterization of β-lactam production by mutant strains 237
V.7 Genetic characterization of the *ccaR* mutants generated by random mutagenesis 238
  V.7.1 Southern blot analysis of NTG mutants 242
  V.7.2 Sequence analysis of the *ccaR* mutants alleles 242
V.8 Complementation of *ccaR* mutants 243
  V.8.1 Complementation of Δ*cob*::*tsr* mutant 243
    V.8.1.1 Characterization of cephamycin C production by complemented Δ*cob*::*tsr* 246
    V.8.1.2 Characterization of production of other β-lactam antibiotics by complemented Δ*cob*::*tsr* 249
  V.8.2 Complementation of Δ*ccaR*::*tsr* mutant with the *ccaR* gene 249
    V.8.2.1 Characterization of cephamycin C production by the complemented Δ*ccaR*::*tsr* mutant 249
  V.8.3 Complementation of the Δ*ccaR*::*tsr* mutant with truncated *ccaR* genes 253
    V.8.3.1 Characterization of cephamycin C production by the Δ*ccaR*::*tsr* mutant complemented with truncated *ccaR* genes 256
  V.8.4 Complementation of the Δ*ccaR*::*tsr* mutant with mutant *ccaR* alleles 259
    V.8.4.1 Characterization of cephamycin C production by the Δ*ccaR*::*tsr* mutant complemented with mutant *ccaR* alleles 259
  V.8.5 Complementation of NTG mutants 262
    V.8.5.1 Characterization of cephamycin C production by the complemented NTG 26 and NTG 32 mutants 265
V.9 Strength of the *lat* promoter in wildtype and *ccaR* mutants 268
  V.9.1 Promoter probe analysis of the *lat* promoter 268
V.10 CcaR-dependent regulation of the *lat* promoter 270
  V.10.1 Characterization of LAT production by the complemented Δ*ccaR*::*tsr* mutant 270
V.11 Time course of production of CcaR during growth of *S. clavuligerus* 270
  V.11.1 Analysis of the time course of CcaR and cephamycin biosynthetic enzyme production by western blot analysis 271
V.12 Attempts to titrate the regulator of antibiotic biosynthesis 271

V.12.1 Characterization of cephamycin C production by *S. clavuligerus* containing high copy number plasmids 274
VI. Discussion 275
VI.I Sequencing of the cephamycin biosynthetic cluster 275
VI.2 Analysis of regulatory genes in the cephamycin cluster 279
VI.3 Analysis of *lat*, *pcbAB* and *pcbC* regulation 292
VI.4 Conclusions 304
VII. Bibliography 306

LIST OF TABLES

Table	Description	Page
1	Promoter probe analysis in *S. lividans* and *S. clavuligerus*	24
2	Bacterial strains list.	42
3	Plasmids list.	62
4	Oligonucleotides list.	80
5	Plasmids and materials used to complete the sequencing of the cosmid fragments	94
6	Differences between published sequence data and new data	115
7	Proteins encoded by the *S. clavuligerus* cephamycin cluster	117
8	Proteins homologous to the new *S. clavuligerus* gene products	118
9	Proteins showing significant homology to the CcaR protein	134
10	Cross hybridization of *S. clavuligerus* cluster genes with other actinomycetes	145
11	Promoter probe analysis of *lat* gene fragments	167
12	Production of cephamycin by the *lat* mutants	185
13	Production of biosynthetic enzymes by the *lat* mutants	190
14	Production of cephamycin by complemented *lat* mutants	200
15	Promoter probe analysis of *lat* promoter in *lat* mutants	207
16	Production of β-lactam antibiotics by *ccaR*, *cob* and NTG mutants	239
17	Production of β-lactam antibiotics by complemented Δ*cob*::*tsr* mutants	250
18	Promoter probe analysis in *ccaR* mutants	269

LIST OF FIGURES

Figure	Description	Page
1	β-lactam metabolites produced by *S. clavuligerus*.	3
2	Partial pathway for the biosynthesis of clavulanic acid and clavams by *S. clavuligerus*.	6
3	Pathway for cephamycin C biosynthesis in *S. clavuligerus*.	11
4	Organization of the *lat-pcbAB* intergenic region.	21
5	Organization of the *pcbAB-pcbC* intergenic region.	23
6	Transcriptional organization of the *lat*, *pcbAB* and the *pcbC* genes.	27
7	Diagrammatic representation of the cephamycin biosynthetic gene cluster in *S. clavuligerus*.	92
8	Compiled sequence data for the cephamycin cluster of *S. clavuligerus*.	96
9	Diagrammatic representation of the cephamycin biosynthetic gene cluster as determined from the compiled sequence data.	114
10	Pathway for α-aminoadipate biosynthesis in *S. clavuligerus*.	124
11	Alignment of proteins homologous to PCD.	127
12	Alignment of BLIP and the Blp protein.	132
13	Alignment of proteins homologous to CcaR.	136
14	Estimation of the correct start codon for CcaR.	139
15	Comparison of the cephamycin biosynthetic clusters from *S. clavuligerus* and *N. lactamdurans*.	142
16	Cross hybridization of the *ccaR* gene with other actinomycetes.	147
17	Reconstruction of the *xylE* reporter gene cassette from pIJ4083 in *E. coli*.	153
18	Construction of the mid-copy number promoter probe vector pDA2500.	155
19	Construction of the single copy promoter probe vector pDA2000.	158
20	*E. coli/Streptomyces* promoter probe shuttle vectors pDA2000 and pDA2500.	160
21	Confirmation of pDA2000-based vector insertion by genomic PCR of transformed strains.	163
22	Map of the *lat* gene-containing DNA and pDA2000 constructs created with fragments of DNA from this region.	165

23 Mutated *lat* genes generated by partial deletion of the *lat* gene via homologous recombination. 171

24 Strategy for the partial deletion of the *lat* gene and replacement with *tsr*. 173

25 Strategy for the partial deletion of the *lat* gene and replacement with *tsr* and the FKMT terminator. 176

26 Strategy for the deletion of an internal fragment of the *lat* gene. 180

27 Confirmation of gene replacement in *lat* mutants by Southern blot analysis. 182

28 Presence of the LAT, IPNS and DAOCS biosynthetic enzymes in the *lat* mutant strains. 188

29 Presence of cephamycin biosynthetic enzymes in the *lat* mutant strains. 193

30 pSET152-based recombinant plasmids containing the *lat* gene. 196

31 Confirmation of pSET152-based vector insertion by genomic PCR of transformed strains. 198

32 Western analysis of *lat* mutants complemented with pSET152-based recombinant plasmids containing the *lat* gene. 203

33 SDS-PAGE analysis of *lat* mutants complemented with pSET152-based recombinant plasmids containing the *lat* gene. 205

34 Diagrammatic representation of the *apr* cassette containing the bifunctional *apr* marker. 212

35 Strategy for insertional inactivation of the *ccaR* gene and replacement with the *apr* cassette. 214

36 Strategy for insertional inactivation of the *orf11* gene and replacement with the *apr* cassette. 217

37 Strategy for the deletion of the *ccaR* gene and replacement with the *tsr* marker. 221

38 Strategy for the deletion of the *cob* genes and replacement with the *tsr* marker. 223

39 Confirmation of gene replacement in *ccaR*, *orf11* and *cob* mutants by Southern blot analysis. 225

40 Presence of the CcaR protein and the cephamycin biosynthetic enzymes in *ccaR* mutant strains as detected by western analysis. 229

41 Presence of the CcaR protein and cephamycin biosynthetic enzymes in the NTG mutant strains as detected by western analysis. 233

42 Presence of the CcaR protein and the cephamycin biosynthetic enzymes in the *ccaR* and NTG mutant strains. 236

43 HPLC analysis of clavulanic acid and clavam metabolite production by cultures grown in SOY fermentation medium. 241

44 Sequence determination of the *ccaR* alleles from *ccaR* mutants. 245

45 Map of the *cob* gene-containing DNA and pSET152-based recombinant plasmids containing the *cob* genes. 248

46 Map of the *ccaR* gene-containing DNA and pSET152-based recombinant plasmids containing the *ccaR* gene. 252

47 Complementation of the Δ*ccaR*::*tsr*A1 mutant with pSET152-based recombinant plasmids containing the *ccaR* gene. 255

48 Complementation of the Δ*ccaR*::*tsr*A1 mutant with pSET152-based recombinant plasmids containing truncated *ccaR* genes. 258

49 Complementation of the Δ*ccaR*::*tsr*A1 mutant with pSET152-based recombinant plasmids containing mutant *ccaR* genes. 261

50 pSET152-based recombinant plasmids containing the *ccaR* or *lat* genes. 264

51 Complementation of the NTG 26 and NTG 32 mutants with pSET152-based recombinant plasmids containing the *ccaR* gene. 267

52 Western blot analysis of an *S. clavuligerus* time course. 273

LIST OF ABBREVIATIONS

αAA	α-aminoadipate
ACT	actinorhodin
ACV	δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine
ACVS	δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase
amp	Ampicillin
apr	Apramycin resistance marker
apr	Apramycin
apr^{R}	Apramycin resistant phenotype
BLAST	Basic local alignment search tool
BLIP	β-lactamase inhibitory protein
C23O	Catechol 2,3 dioxygenase
cob	*ccaR*, *orf11* and *blp* group of genes
CAS	Clavaminate synthase
DAC	Desacetylcephalosporin C
DACS	Desacetylcephalosporin C synthase
DAOCS	Desacetoxycephalosporin C synthase
dNTPs	Deoxynucleoside triphosphates
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ExoIII	Exonuclease III
FKMT	31-demethyl-FK506-*O*-methyltransferase
gBLAST	Gapped basic local alignment search tool
GST	Glutathione-S-transferase
HPLC	High performance liquid chromatography
HTH	Helix-turn-helix
hyg	Hygromycin
hyg^{R}	Hygromycin resistant phenotype
IPNE	Isopenicillin N epimerase
IPNS	Isopenicillin N synthase
IPTG	Isopropyl-β-D-thiogalactopyranoside
ISP3	ISP medium 3
Kan	Kanamycin
KOAc	Potassium acetate

LAT	Lysine ε-aminotransferase
MBP	Maltose binding protein
MCS	Multiple cloning site
MOPS	3-[N-Morpholino]propanesulfonic acid
MSP	Maximal segment pair
NH_4OAc	Ammonium acetate
NaOAc	Sodium acetate
neo	Neomycin
neo	Neomycin resistance gene
NTG	N'-methyl-N'-nitro-N-nitrosoguanidine
OCDAC	*O*-carbamoyl-desacetylcephalosporin C
ori	origin of replication
ORF	Open reading frame
RED	undecylprodigiosin
PAG	Polyacrylamide gel
PAH	Proclavaminate amidinohydrolase
PBP	Penicillin binding protein
PCD	Piperideine-6-carboxylate dehydrogenase
PEG	Polyethylene glycol
PSI-BLAST	Position specific iterated BLAST
SA	Starch-asparagine media
S/D	Shine-Dalgarno sequence
SDS	Sodium dodecyl sulfate
TBST	Tris-buffered saline plus tween
TBSTM	Tris-buffered saline plus tween plus skim milk
Tc	Tetracycline
TES	N-tris[hydroxymethyl]methyl-2-aminoethanesulfonic acid
Tris	Tris(hydroxylmethyl)-aminomethane
TSBM	Trypticase soy broth plus 1% (w/v) maltose
TSBS	Trypticase soy broth plus 1% (w/v) starch
TSF	Translational stop fragment
tsp	Transcription start point
tsr	Thiostrepton resistance marker
tsr	Thiostrepton
tsr^R	Thiostrepton resistant phenotype
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside

I. Introduction

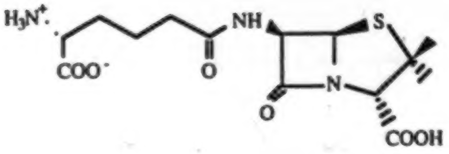
I.1 β-lactam antibiotics

β-lactam antibiotics account for more than 50% of all antibiotics prescribed worldwide. The popularity of these antibiotics arises from their mechanism of action, as β-lactam antibiotics target bacterial cell wall biosynthesis and therefore have low toxicity to humans. These antibiotics prevent peptidoglycan maturation by specifically inhibiting the transpeptidase cross-linking reaction as the β-lactam antibiotics have a structural similarity to the D-alanyl-D-alanine substrate of the cross-linking reaction (130). Decreased peptide cross-links result in the production of a defective peptidoglycan backbone of the cell wall which will cause the death of growing bacteria. Many proteins in the cell membrane have been identified which bind β-lactam antibiotics with differing degrees of affinity. These proteins are referred to as penicillin binding proteins (PBP) and they play roles within cell wall assembly.

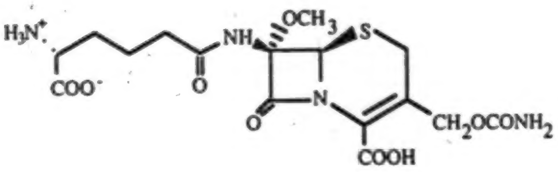
The fungus *Penicillium notatum* first demonstrated the ability to produce penicillin (76), an antibiotic containing the 5-membered sulfur-containing penam ring structure, while the 6-membered sulfur-containing cephem ring structure found in the cephalosporins was first identified after production by *Cephalosporium acremonium* (157). *Streptomyces* and *Nocardia* spp. have demonstrated the ability to produce cephamycins, which are closely related to cephalosporins in that they contain the 6-membered sulfur-containing cephem ring structure except that a methoxy group is added at carbon 7 (130). This allows the cephamycins to have greater intrinsic resistance to β-lactamases, which cleaves the β-lactam ring of the penicillins or cephalosporins. The clavams differ from the penicillin/cephalosporin β-lactam antibiotics in that they have a 5-membered oxygen-containing oxypenam ring structure.

Streptomyces clavuligerus produces at least 21 different secondary metabolites (96) including two major groups of β-lactam metabolites (Fig. 1); the sulfur-containing metabolites penicillin N and cephamycin C as well as the oxygen-containing metabolites which include clavulanic acid and several structurally related non-clavulanic acid clavams. The non-clavulanic acid clavams differ from clavulanic acid in the stereochemistry of the clavam nucleus and the nature of the substituent groups (154) and for the remainder of this study the group of metabolites will be referred to as only clavams. This ring stereochemistry also makes clavulanic acid a weak antibiotic but a potent irreversible β-lactamase inhibitor (186). Clavulanic acid has also become a very important biopharmaceutical because of its potent β-lactamase inhibitory activity. The combination of clavulanic acid and the penicillin derivatives amoxycillin or ticarcillin are sold as the

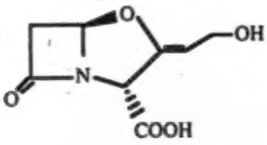
Figure 1. β-lactam metabolites produced by *S. clavuligerus*.


H3N+
COO-
O
NH
S
N
O
COOH

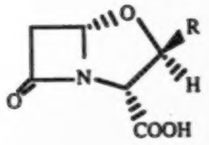
penicillin N


H3N+
COO-
O
NH
OCH3
S
N
O
CH2OCONH2
COOH

cephamycin C



clavulanic acid



clavams

R = COOH	**clavam-2-carboxylate**
CH_2OH	**2-hydroxymethylclavam**
CH_2OCHO	**2-formyloxymethylclavam**
CH_2CHNH_2COOH	**alanylclavam**

chemotherapeutics Augmentin™ or Timentin™ (12). The clavams do not possess any β-lactamase inhibitory activity but they do have antibacterial or antifungal activities.

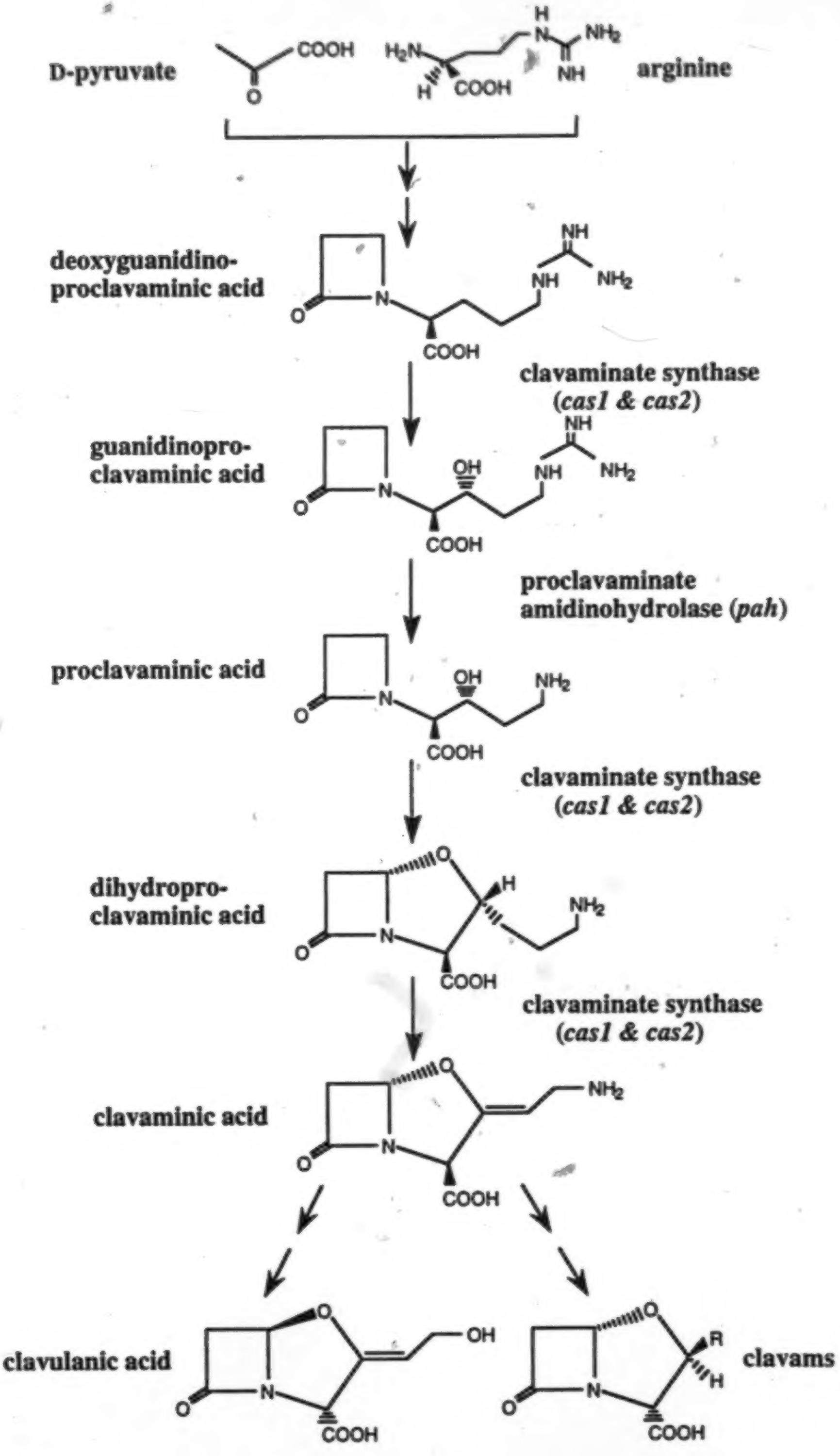
I.2 Clavulanic acid/clavam biosynthetic pathways and clusters

The clavulanic acid and clavam biosynthetic pathways are only beginning to be investigated and are understood in less detail than the cephamycin pathway. Rather than biochemically determining each of the pathway steps the biosynthetic pathways are being characterized primarily through the genetic analysis and sequencing of the biosynthetic clusters. The activities and/or roles of the genes in antibiotic biosynthesis are primarily being estimated via database homology searches and gene disruption experiments but they have included some characterization of purified proteins and/or expressed genes.

It has been determined that clavulanic acid biosynthesis (Fig. 2) begins with the condensation of arginine and D-pyruvate (211, 218). After an as yet undetermined number of steps, the intermediate guanidinoproclavaminic acid is hydrolyzed to convert the guanidino group to an amino group and yield proclavaminic acid by the action of an proclavaminate amidinohydrolase (PAH) (68). The PAH protein has been purified from extracts of *S. clavuligerus* and N-terminal sequencing of the PAH protein confirmed that the *pah* gene [(95); also referred to as *cla* (6)] coded for PAH and was located directly adjacent to the gene for the clavaminate synthase (CAS) in the clavulanic acid biosynthetic cluster.

The CAS enzyme was the first clavulanic acid biosynthetic activity detected in *S. clavuligerus* cell-extracts and was demonstrated to catalyze the conversion of proclavaminic acid to clavaminic acid (69). CAS enzyme activity was also found in cell-extracts generated from the clavulanic acid producers *S. jumonjinensis* and *S. katsurahamanus* and the cephamycin producer *S. lipmanii.* CAS was determined to catalyze two successive reactions; the oxidative cyclization of proclavaminic acid to the intermediate dihydroproclavaminic acid followed by a desaturation to clavaminic acid (15, 191, 192). Attempts to purify the enzyme led to the discovery of two nearly identical isozymes of the CAS protein referred to as CAS1 and CAS2 with identical enzyme activities and cofactor requirements (192). Each of the proteins is encoded by a separate gene called *cas1* and *cas2* which are 87% identical as well as showing homology to the oxygenases of the penicillin/cephamycin biosynthetic pathway (145). Baldwin et al. (17) demonstrated that the CAS enzymes also catalyze the hydroxylation of deoxyguanidinoproclavaminic acid to guanidinoproclavaminic acid. *S. antibioticus* Tü 1718 makes clavams (valclavam and 2-hydroxyethyl clavam) but does not produce clavulanic acid and it was possible to detect PAH and CAS enzyme activities in cell-

Figure 2. Partial pathway for the biosynthesis of clavulanic acid and clavams by *S. clavuligerus*.


D-pyruvate
arginine
deoxyguanidino-proclavaminic acid
clavaminate synthase (cas1 & cas2)
guanidinopro-clavaminic acid
proclavaminate amidinohydrolase (pah)
proclavaminic acid
clavaminate synthase (cas1 & cas2)
dihydropro-clavaminic acid
clavaminate synthase (cas1 & cas2)
clavaminic acid
clavulanic acid
clavams

extracts (16). Janc et al. (103) was able to purify a single CAS enzyme from these *S. antibioticus* cell-extracts and this enzyme called CAS3 also possessed all three enzymatic activities.

Heterologous expression of the *cas1* and *cas2* genes in *E. coli* confirmed that the CAS1 and CAS2 proteins are multifunctional oxygenases which could catalyze all three reactions (35). Wu et al. (229) demonstrated that the addition of heterologously produced CAS2 and PAH proteins could convert deoxyguanidinoproclavaminic acid to clavaminic acid *in vitro*. If the PAH enzyme was excluded from the reaction the CAS mediated conversion was arrested at guanidinoproclavaminic acid.

The duplication of the *cas* gene may be an attempt by the organism to address the fact that the CAS step appears to be the rate limiting step as it has been demonstrated that proclavaminic acid accumulates within the cells during the fermentation (69, 192). The bottleneck may be the result of the oxidative self-inactivation of the CAS enzyme during catalysis (192). However, Egan et al. (66) has demonstrated that both the clavulanic acid and clavam biosynthetic pathways can share intermediates up to and including clavaminic acid suggesting that maybe more of the early genes are duplicated like *cas*.

An undetermined number of "late" steps are required to convert clavaminic acid to clavulanic acid. One of these steps appears to be the reduction of an unstable aldehyde intermediate to clavulanic acid and this step is catalyzed by the clavulanate aldehyde dehydrogenase (158).

The clavulanic acid biosynthetic cluster is located directly adjacent to the cephamycin cluster in *S. clavuligerus*, *S. jumonjinensis* and *S. katsuharamanus* to form a biosynthetic "super-cluster" (222) and a large portion of the *S. clavuligerus* cluster has been sequenced (95, 107). The *cas1* gene is apparently not located near the cephamycin/clavulanic acid supercluster since the *cas1* gene is separated from the *cas2* gene by greater than 40 kb (151).

Mutants of the *cas2* gene were created by gene disruption via homologous recombination and the resulting *cas2* mutant strain was unable to produce clavulanic acid in starch-asparagine (SA) medium (168). The *cas2* mutant had an intact *cas1* gene but was unable to produce clavulanic acid which suggested that the two *cas* genes may be differentially regulated. When the *cas2* mutant was grown in complex SOY medium, which was used during the purification of both CAS isoenzymes, the *cas2* mutant was able to produce 60% of wildtype clavulanic acid levels. Northern analysis of the *cas1* and *cas2* genes demonstrated *cas2* expression in both SA and SOY media while *cas1* expression was restricted to the SOY medium. The *cas1* monocistronic transcript was 1.3 kb while the *cas2* coding region was present as a 1.2 kb monocistronic transcript and a

5.3 kb polycistronic transcript believed to contain *orf2*, *orf3*, *pah* and *cas2*. Therefore the *cas2* mutants show a leaky phenotype in SOY medium as the *cas1* gene compensates for the mutation but in SA medium the *cas2* gene is essential for clavulanic acid production. Clavam production in *S. clavuligerus* occurs exclusively in SOY medium as only clavulanic acid is produced in SA medium. Since wildtype *S. clavuligerus* generates clavaminic acid in SA medium but not alanylclavam some of the other clavam-specific biosynthetic genes in addition to *cas1* must also be under this transcriptional control.

Conversely, mutants of the *cas1* gene were created by gene disruption and the resulting *cas1* mutant strain produced wildtype levels of clavulanic acid in SA medium but the levels of clavulanic acid and the majority of the clavams were much less in SOY medium (152). Therefore the *cas1* mutant is compensated somewhat by the *cas2* gene to allow reduced production of the clavams and clavulanic acid. This suggested that the CAS1 and CAS2 proteins contribute equally to a common intracellular pool of clavaminic acid which can be drawn on by the clavam or clavulanic acid biosynthetic pathways. This supports the conclusions of Egan et al. (66) which suggested that both biosynthetic pathways share intermediates up to the point of clavaminic acid. A *cas1/cas2* double mutant strain created by gene disruption was unable to produce clavulanic acid or clavams in either SOY or SA media (152).

The ability of the duplicated biosynthetic genes to compensate for a mutation in the homologous clavulanic acid biosynthetic gene is not restricted to the *cas1/cas2* genes (152). Mutation of the *pah* gene or the *orf6* gene (encodes a protein homologous to an ornithine transacetylase) results in a clavulanic acid non-producing phenotype in SA medium but growth of either of these mutant strains in SOY medium returns clavulanic acid production. Homologs for these genes have not been found but they must exist and they are also transcriptionally regulated in a similar manner to *cas1*.

This however is not true for the *orf2* gene which encodes a protein homologous to acetohydroxyacid synthase which is involved in branched side-chain amino acid biosynthesis (107). Since valine is a precursor of penicillin/cephamycin biosynthesis it was not clear at the onset whether the *orf2* gene product was involved in clavulanic acid or cephamycin biosynthesis. Mutants of the *orf2* gene were created by gene disruption and the resulting *orf2* mutant strains were unable to produce clavulanic acid in SA medium but produced cephamycin (67). Analysis of *orf2* mutant culture supernatants failed to demonstrate any imidazole-reacting metabolites suggesting that the *orf2* gene product is involved in one of the earliest steps of clavulanic acid biosynthesis. Unlike some of the other "early gene" mutations, the *orf2* mutants retained this clavulanic acid negative phenotype in SOY medium and unexpectedly, the *orf2* mutants were also unable to make clavams. The clavulanic acid and clavam negative phenotypes suggest that the *orf2* gene does not have a functional homolog which can compensate for the *orf2* mutation. *S. antibioticus* Tü 1718 which does not produce clavulanic acid (16), must contain an *orf2* like-gene in the clavam biosynthetic cluster to allow for the production of clavams.

Located within the clavulanic acid biosynthetic cluster is the *claR* gene and the putative ClaR protein shows homology to the LysR family of transcriptional activators. This homology is primarily concentrated around the helix-turn-helix domain within the protein (165). Mutants of the *claR* gene were created by gene disruption and the resulting *claR* mutant strain was unable to produce clavulanic acid in SA or SOY media. The *claR* mutant however accumulates clavaminic acid which suggested that *claR* is responsible for the transcriptional regulation of the "late genes" which would catalyze the reactions after the clavaminic acid intermediate is made. Northern analysis demonstrated that the *claR* mutation did not affect the accumulation of the *orf2/orf3/pah/cas2* polycistronic transcript or the *cas2* monocistronic transcript. High resolution S1 nuclease transcript analysis determined that *orf6* was not regulated by *claR* but transcripts containing *orf7*, *orf9* and *orf10* were transcriptionally regulated.

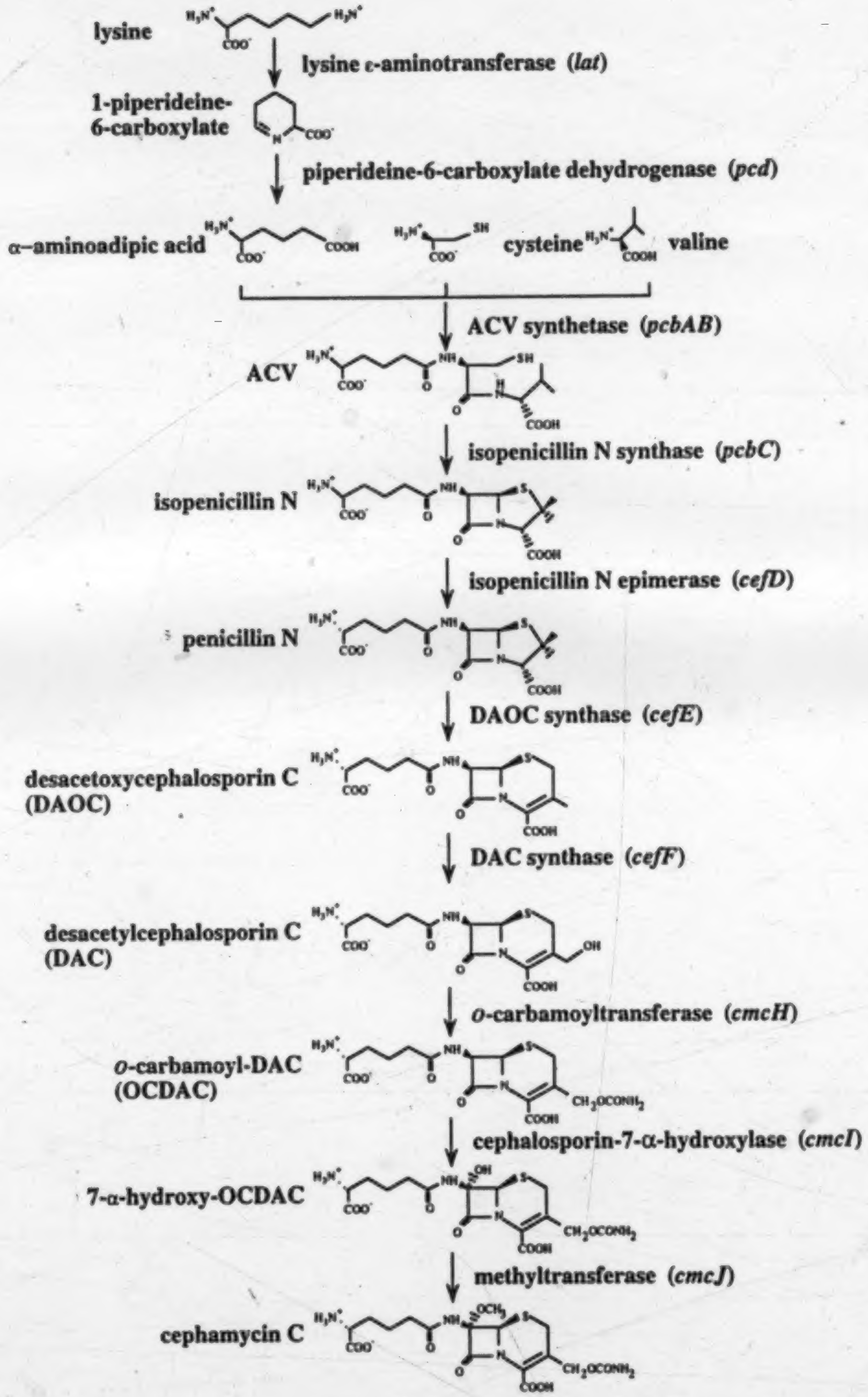
I.3 Penicillin/cephalosporin/cephamycin biosynthetic pathway

Unlike clavulanic acid and clavam biosynthesis the penicillin/cephamycin biosynthetic pathway was characterized at a time when molecular biology was in its infancy and therefore enzymology and biochemistry rather than cloning and sequencing were the tools used to analyze the pathway. As a result, the majority of the reactions and enzymes of the biosynthetic pathway are understood in great detail (Fig. 3).

Penicillin/cephalosporin biosynthesis occurs in both eukaryotic filamentous fungi (*Penicillium*, *Cephalosporium* and *Aspergillus*), filamentous bacteria (*Streptomyces* and *Nocardia*) and other non-filamentous bacteria (*Flavobacterium* and *Lysobacter*) (170). Producing organisms, irrespective of their fungal or bacterial origins, contain a core of conserved enzyme activities which lead to antibiotic production via a common pathway but significant differences do exist between these systems. Emphasis has been placed on the initial enzymes of the pathway and particularly what information is known about them in the organism of interest, *S. clavuligerus*.

In the fungal systems α–aminoadipate (αAA), one of the precursor amino acids, is produced as an intermediate of lysine biosynthesis and can be drawn off from primary metabolism for penicillin/cephalosporin biosynthesis (170). Lysine biosynthesis in

Figure 3. Pathway for cephamycin C biosynthesis in *S. clavuligerus*.


lysine
lysine ε-aminotransferase (*lat*)
1-piperideine-6-carboxylate
piperideine-6-carboxylate dehydrogenase (*pcd*)
α–aminoadipic acid
cysteine
valine
ACV synthetase (*pcbAB*)
ACV
isopenicillin N synthase (*pcbC*)
isopenicillin N
isopenicillin N epimerase (*cefD*)
penicillin N
DAOC synthase (*cefE*)
desacetoxycephalosporin C (DAOC)
DAC synthase (*cefF*)
desacetylcephalosporin C (DAC)
o-carbamoyltransferase (*cmcH*)
o-carbamoyl-DAC (OCDAC)
cephalosporin-7-α-hydroxylase (*cmcI*)
7-α-hydroxy-OCDAC
methyltransferase (*cmcJ*)
cephamycin C

bacteria proceeds via the diaminopimelic acid pathway of which αAA is not an intermediate.

Whitney et al. (227) demonstrated that radiolabeled lysine becomes incorporated into cephamycin C in *S. clavuligerus*. The other precursor amino acids valine and cysteine are primary metabolic amino acids but it was not known how lysine was converted to αAA so that it could be incorporated into the penicillin structure. Madduri et al. (141) demonstrated that two pathways for lysine catabolism existed in *S. clavuligerus* initiating with lysine ε-aminotransferase (LAT) or cadaverine aminotransferase. The cadaverine aminotransferase activity was present in all nine *Streptomyces* spp. (2 β-lactam producers and 7 non-β-lactam producers) and the one *Nocardia* spp. but LAT activity was present in only the three penicillin/cephamycin producers. It was concluded that the cadaverine pathway present in all species was necessary for utilizing lysine as a nitrogen source, while the LAT pathway was specific for secondary metabolism. Since LAT activity is only present in penicillin/cephamycin producers it is considered to be the first step of antibiotic biosynthesis. Kern et al. (119) had previously demonstrated that *Nocardia lactamdurans* contained the LAT enzyme which converted lysine to 1-piperideine-6-carboxylate as the first of a two step reaction converting lysine to αAA.

Accumulation of αAA was demonstrated during growth and the precursor was later drawn off for cephamycin production (139). In *S. clavuligerus*, LAT activity peaks during the end of exponential growth and decreases after antibiotic production occurs (141, 188). Romero et al. (188) determined that the LAT activity was very unstable in frozen cell-extracts and therefore LAT activity was immediately assayed in fresh cell-extracts. LAT enzyme activity required the addition of pyridoxal phosphate and α-ketoglutarate as cofactors.

The product of LAT action, 1-piperideine-6-carboxylate, is presumed to require the activity of a NAD-dependent dehydrogenase to be converted into αAA (119). Recently, a protein with piperideine-6-carboxylate dehydrogenase activity has been purified from *S. clavuligerus* and this activity like that of the LAT protein was specific to penicillin/cephamycin producers (55). Time course analysis suggests that piperideine-6-carboxylate dehydrogenase activity correlates with cephamycin C production but peaks later than LAT activity and is more similar to some of the enzymes from the middle of the cephamycin pathway.

The δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase (ACVS) enzyme catalyzes the condensation of the three precursor amino acids, αAA, valine and cysteine into a tripeptide called δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV). Cell free synthesis of the ACV tripeptide was demonstrated in cell-extracts of *S. clavuligerus* when the appropriate component amino acids were added along with ATP and an ATP regenerating system involving phosphoenol pyruvate and pyruvate kinase (116). Purified ACVS was also capable of catalyzing the ligation of the three precursor amino acids (117, 235) and the estimated size of the purified enzyme was between 500-560 KD (196).

Synthesis of the ACV tripeptide occurs via the nonribosomal protein thiotemplate mechanism (121). Characteristic of antibiotic peptide synthetases, the multifunctional ACVS protein contains three repeated domains and each domain contains the active site for the addition of αAA, cysteine and valine, respectively. Each domain is responsible for the ATP-mediated activation of the amino acids into aminoacyl adenylates which bind to the phosphopantetheine cofactor by thioester linkages and the amino acids become joined in a sequential manner by transpeptidation. The valine domain is also responsible for the epimerization of L-valine to D-valine, while a final domain near the C-terminal end of the protein which resembles thioesterases is thought to release the tripeptide from the active site of the protein.

The ACV tripeptide undergoes an oxidative cyclization by the isopenicillin N synthase (IPNS) enzyme to form the β-lactam and thiazolidine rings (170) and thereby generates the first bioactive molecule. The IPNS catalyzed reaction requires the addition of molecular oxygen, dithiothreitol (DTT), ferrous ions and ascorbic acid as cofactors (113). *S. clavuligerus* cell-extracts and purified IPNS protein have been demonstrated to convert ACV to isopenicillin N (110). The DTT is believed to serve two functions as it prevents inactivation of the essential thiol groups that may play a role in catalysis or maintenance of the proper protein conformation as well as maintaining bis-ACV as the monomeric natural substrate.

The fungi which produce only penicillin metabolites rely on the acylcoenzymeA-isopenicillin acyltransferase complex to remove the αAA side chain from isopenicillin N to give 6-aminopenicillanic acid (18, 226). The side chain is replaced with a range of substituted acetic acid non-polar side chains like phenylacetyl-coenzymeA to produce a variety of penicillins.

In the cephalosporin/cephamycin producing organisms the first step that commits the pathway to the production of cephalosporin/cephamycin is catalyzed by isopenicillin N epimerase (IPNE) as cell-extracts from *Cephalosporium* contain IPNE activity but *Penicillium* cell-extracts do not (105). IPNE activity epimerizes the L-form of the αAA side chain of isopenicillin N to the D-form to give penicillin N. Cell-extracts of *S. clavuligerus* contain IPNE activity (114) and the IPNE protein has been purified from these cell-extracts (217). IPNE activity does not require the addition of any cofactors for maximal activity but IPNE protein purified in the absence of pyridoxal phosphate becomes irreversibly inactivated (114).

The oxidative expansion of the 5-membered thiazolidine ring of penicillin N into a 6-membered dihydrothiazine ring characteristic of all cephalosporins is catalyzed by the desacetoxycephalosporin C synthase (DAOCS). DAOCS activity was first detected in *C. acremonium* cell-extracts and required ferrous ions, molecular oxygen, DTT, ascorbic acid and α-ketoglutarate for cofactors (128). Similar cofactors were required for the detection of DAOCS activity in *S. clavuligerus* cell-extracts (115).

The hydroxylation of the desacetoxycephalosporin C to desacetylcephalosporin C (DAC) is catalyzed by the DAC synthase (DACS) enzyme (14). Detection of DACS enzyme activity requires the addition of molecular oxygen, DTT, ferrous ions and ascorbic acid as cofactors to both *C. acremonium* (64) and *S. clavuligerus* cell-extracts (115). Attempts to separate DAOCS and DACS activities in *C. acremonium* cell-extracts were unsuccessful and it was demonstrated that both activities arose from a single bifunctional protein (64). The DAOCS and DACS activities were demonstrated to be present on two separate proteins in *S. clavuligerus* (115).

The fungus *C. acremonium* produces cephalosporin C, rather than cephamycin like *S. clavuligerus*, and therefore the pathways diverge at DAC. DAC is acetylated to cephalosporin C by the DAC acetyltransferase (82).

Cephamycin C production in *S. clavuligerus* requires three steps to convert DAC to cephamycin C. The addition of a carbamoyl group from carbamoyl phosphate to DAC is catalyzed by the *O*-carbamoyltransferase which results in the synthesis of *O*-carbamoyl-DAC (OCDAC) (31). OCDAC is converted to 7-α-hydroxy-OCDAC by the cephalosporin-7-α-hydroxylase which has been purified from *S. clavuligerus* (231). The methyltransferase catalyzes the transfer of the methyl group from S-adenosylmethionine to 7-α-hydroxy-OCDAC to synthesize cephamycin C (170). Enguita et al. (70) have demonstrated that the cephalosporin-7-α-hydroxylase and the methyltransferase form a heterodimer *in vivo*.

I.4 Cephamycin biosynthetic cluster

The genes responsible for antibiotic biosynthesis, transport, resistance and regulation are organized into clusters in antibiotic producing *Streptomyces* (146). For *S. clavuligerus* the majority of the cephamycin biosynthetic enzymes were biochemically characterized before the genes coding for the enzymes were discovered and sequenced. These genes were often discovered during experiments initiated with N-terminal amino acid sequence data obtained from the purified biosynthetic enzyme. Conversely, the majority of the genes from the *N. lactamdurans* cephamycin cluster were determined through a large scale sequencing project which completed the sequence of what is felt to be the entire biosynthetic cluster.

The genes encoding three of the earliest enzymes in the biosynthetic pathway; LAT, ACVS and IPNS are designated *lat*, *pcbAB* and *pcbC* (108) and are referred to as the "early" genes. The LAT protein catalyzes the first of a two step reaction as lysine is converted to 1-piperideine-6-carboxylate which is then converted to αAA by the action of a dehydrogenase (119). No candidate genes encoding a dehydrogenase have been discovered in any bacterial cephamycin clusters studied to date (170) which led to speculation that the reaction might be catalyzed by a primary metabolic enzyme.

The *lat* gene from *S. clavuligerus* was discovered and sequenced after finding a fragment of DNA which was capable of heterologous expression of LAT activity in the penicillin/cephalosporin non-producer *S. lividans* (142, 215) and was also re-sequenced in a different study (234). The *lat* gene from *N. lactamdurans* has also been sequenced (50). Southern blot analyses demonstrated that genes homologous to the *N. lactamdurans* or *S. clavuligerus lat* genes were present only in penicillin/cephamycin producing organisms (50, 215), like the presence of LAT activity.

In both *N. lactamdurans* or *S. clavuligerus* the *lat* gene is located directly adjacent to the *pcbAB* gene coding for the ACVS protein. The entire *pcbAB* gene from *N. lactamdurans* has been sequenced and the gene requires almost 11 kb of DNA to code for a protein with a deduced molecular weight of 404 KD (52). The multidomain nature of the *pcbAB* gene could be observed as the domains share extensive sequence homology with one another and with domains from other peptide synthetases.

The *S. clavuligerus pcbAB* gene has only been partially sequenced from both the 5' end (215, 234) and the 3' end (63), which would leave approximately 7.8 kb of internal DNA unsequenced, assuming that the *S. clavuligerus* gene is the same size as the gene from *N. lactamdurans*. The *N. lactamdurans* cluster has a *lat-pcbAB* intergenic region of 65 bp (50) while *S. clavuligerus* has a larger intergenic region of 152 bp (234).

The 3' end of the *pcbAB* genes from *N. lactamdurans* and *S. clavuligerus* are directly adjacent to the *pcbC* gene coding for the IPNS protein. The *pcbC* gene from *S. clavuligerus* was sequenced by Leskiw et al. (134) and the 5' and 3' intergenic regions were sequenced by Doran et al. (63) and the *pcbAB* and *pcbC* genes are separated by 31 bp of intergenic DNA. The *pcbC* gene from *N. lactamdurans* was sequenced along with the *pcbAB* gene and the *pcbAB* stop codon is directly adjacent to the *pcbC* start codon (52). The *pcbC* gene was the first penicillin/cephalosporin biosynthetic gene cloned and sequenced and the gene has been sequenced from a number of bacterial and fungal hosts.

All of the *pcbC* gene sequences have a high degree of similarity at the nucleotide and amino acid levels. The lack of introns within the coding region of the fungal genes and the relatively high GC content of the *pcbC* genes suggested that the β-lactam genes originated in a *Streptomyces* spp. and then were transferred to the eukaryotic penicillin/cephalosporin producers (3).

In the *N. lactamdurans* or *S. clavuligerus* cephamycin biosynthetic gene clusters the *lat*, *pcbAB* and *pcbC* genes are ordered sequentially with little or no intergenic spacing between the genes. This close spacing and same transcriptional orientation suggests an operon-like structure may exist which could coordinate expression of the three genes (170). The transcriptional organization of these genes has been determined by northern analysis, primer extension and/or S1 nuclease mapping and the *lat* promoter is thought to direct the synthesis of a polycistronic transcript of approximately 14 kb responsible for expression of *lat*, *pcbAB* and *pcbC* (176, 178). The *pcbC* gene from *S. griseus* has been sequenced (84) and the upstream sequence shows homology to the 3' end of the *S. clavuligerus pcbAB* gene suggesting that a similar arrangement may also exist in *S. griseus*.

The *pcbAB* and *pcbC* genes from *C. acremonium* and *P. chrysogenum* are also adjacent to one another but they are separated by a large intergenic region and the genes are divergently expressed to give transcripts of 11 kb and 1.2 kb in size (61, 90). The *penDE* gene responsible for the acylcoenzymeA-isopenicillin acyltransferase is located downstream of the *pcbC* gene in *P. chrysogenum* (61) but is not present in the cephalosporin producer *C. acremonium*.

The genes encoding the three enzymatic steps that create and modify the cephem nucleus in the biosynthetic pathway; IPNE, DAOCS and DACS are designated *cefD*, *cefE* and *cefF* (108) and are referred to as the "middle" genes. The *cefD* and *cefE* genes are directly adjacent to each other in both *N. lactamdurans* and *S. clavuligerus* (53, 123, 124). The *cefD* and *cefE* genes have the same transcriptional orientation and are also tightly linked with an 81 bp and no intergenic region for *S. clavuligerus* and *N. lactamdurans*, respectively. Northern analysis of *S. clavuligerus* RNA detected a *cefD*- and *cefE*-containing transcript of greater than 10 kb (123), which would suggest that additional downstream genes are included with *cefD* and *cefE* within this transcript. The *S. clavuligerus cefF* gene is not included in the *cefD*/*cefE* polycistronic message because the *cefF* gene is located approximately 2.5 kb upstream of the *cefD* gene and is transcribed in the other orientation (122). The *N. lactamdurans cefF* gene is also transcribed in the opposite orientation to *cefD* and *cefE* and is tightly organized with the "late" genes (48). The *cefE* and *cefF* genes share a significant degree of identity at the nucleotide and amino acid level and are believed to have arisen from a gene duplication (122).

Attempts to separate DAOCS and DACS activities in *C. acremonium* cell-extracts were unsuccessful which suggested that both activities arose from a single bifunctional protein (64) and it was determined that this bifunctional protein is encoded by a single gene called *cefEF* (195). The *cefEF* gene may have been generated by a gene fusion as the gene shows significant homology to both the *cefE* and *cefF* genes from *S. clavuligerus*. Alternatively, it has also been proposed that the duplication of the *cefE* and *cefF* genes occurred after the horizontal transfer to the fungal host (197). The *cefD* gene encoding the IPNE protein has not been found in *C. acremonium*.

The *cefG* gene which encodes the acetylcoenzyme-DAC acetyltransferase has been located next to the *cefEF* gene in *C. acremonium* but on a different chromosome than the *pcbAB* and *pcbC* genes (91). The *cefG* and the *penDE* genes which do not have prokaryotic homologs possess multiple introns (91, 214) but the genes which are of prokaryotic origin do not.

The genes encoding the final three enzymatic steps converting DAC to cephamycin C in the biosynthetic pathway; *O*-carbamoyltransferase, cephalosporin-7-α-hydroxylase and methyltransferase, are designated *cmcH*, *cmcI* and *cmcJ* (108) and are referred to as the "late" genes. The *N. lactamdurans cmcH* gene has been sequenced and was located downstream of the *cefF* gene within the cluster (54). As well, a partial *cmcH* sequence was obtained from this same *cefF* downstream region of *S. clavuligerus*. The *cmcI* and *cmcJ* genes from *N. lactamdurans* have been sequenced and they are tightly clustered upstream of the *cefF* and *cmcH* genes and may also be part of an operon (49). The *cmcI* gene from *S. clavuligerus* has been sequenced but has not been placed within the cluster relative to any of the other genes (230) and the *cmcJ* has not yet been sequenced.

The *N. lactamdurans* cluster also contains the *cmcT* gene which encodes a potential transmembrane protein that may function as a penicillin/cephamycin export protein (51). The putative CmcT protein contains 11 transmembrane domains as over 50% of the amino acids are hydrophobic. Overexpression of *cmcT* in *S. lividans* led to the accumulation of the CmcT protein in the cell membrane.

The *N. lactamdurans* cluster appears to be flanked by resistance genes (51). One end of the cluster has the *bla* gene which encodes a β-lactamase while the other end has the *pbp* gene which encodes a PBP. The cell wall bound β-lactamase was shown to be active versus penicillins but the enzyme did not inactivate cephalosporins or cephamycin C. A *bla* gene has also been found within *S. clavuligerus* (173) and this gene may represent one end of the *S. clavuligerus* cluster. The *N. lactamdurans* PBP shows homology to DD-carboxypeptidases and low molecular weight PBPs and was shown to be able bind labeled penicillin in a standard PBP assay but no affinity was observed for cephamycin C. The *pcbR* gene of *S. clavuligerus* is located directly adjacent to the clavulanic acid cluster and represents one end of the cephamycin cluster (166). The *pcbR* gene shows homology to high molecular weight PBPs and *pcbR* mutants have increased sensitivity to penicillins and cephalosporins.

Promoter probe analysis demonstrated very strong activity for the *lat* promoter fragment from *S. clavuligerus* when analyzed in *S. clavuligerus* but not when analyzed in the closely related *S. lividans* (176). The species-specific activation of the *lat* promoter suggested the involvement of a transcriptional activator present only within *S. clavuligerus*. It was felt that a transcriptional activator responsible for regulation of the cephamycin biosynthesis, if one existed, would most likely be located within the cephamycin gene cluster. Since no such gene had yet been described it seemed possible that it might be located in one of the many gaps existing within the sequence of the cluster. The sequence and gene organization of the *N. lactamdurans* cephamycin biosynthetic cluster has been determined but the cluster lacked an apparent transcriptional regulator. A large scale sequencing project was initiated to sequence these gaps and complete the sequence of the *S. clavuligerus* cephamycin cluster.

After the sequencing project had been initiated, the DNA sequences of the transcriptional activator *dclX* (221) or *ccaR* (174) were published. Although a transcriptional activator had been located the sequencing project was continued with the goal of finding all of the genes responsible for the regulation of cephamycin biosynthesis.

I.5 Organization and regulation of the *lat*, *pcbAB* and *pcbC* genes

In *S. clavuligerus* the *lat*, *pcbAB* and *pcbC* genes are ordered sequentially within the cluster with little intergenic spacing between the genes and with the same transcriptional orientation which suggests an operon-like structure that would coordinate expression of the three genes (170). Northern analysis of RNA isolated from wildtype *S. clavuligerus* could not demonstrate the presence of monocistronic *lat* or *pcbAB* transcripts but rather smears of high molecular weight mRNA were detected with both *lat*- and *pcbAB*-specific probes (176). The northern analysis however did demonstrate the presence of a monocistronic *pcbC* transcript. The absence of a monocistronic *lat* message suggested that *lat* and *pcbAB* were present on the same transcript and *pcbC* was present on its own monocistronic transcript.

Sequence analysis of the *lat-pcbAB* intergenic region suggested that a high GC stem-loop structure which may function as a transcription terminator was present between the genes (215, 234). A potential anti-terminator box was found upstream of the stem-loop structure and this region was postulated to allow conditional regulation of transcription termination at the stem-loop (Fig. 4) (234). The *lat* stem-loop structure must allow transcription to proceed into the *pcbAB* gene since a promoter region has not been detected after the stem-loop structure by promoter probe analysis or S1 nuclease transcript analysis (176). Northern blot analysis did not detect a monocistronic *lat* transcript but rather an accumulation of degraded mRNA decreasing from a very high molecular weight.

Ironically, the *pcbAB-pcbC* intergenic region did not contain any strong stem-loop structures which could function as a transcription terminator (63) yet smears of high molecular weight mRNA that would be evidence of a *pcbC*-containing polycistronic message were not detected with a *pcbC*-specific probe (176). Only the monocistronic *pcbC* transcript was detected A monocistronic *pcbC* transcript has also been detected in *S. griseus* (84). The 31 bp *pcbAB-pcbC* intergenic region does contain regions of dyad symmetry capable of forming a GC-rich stem-loop structure but this stem-loop structure was considered to be too small to successfully terminate transcription (Fig. 5) (63, 109).

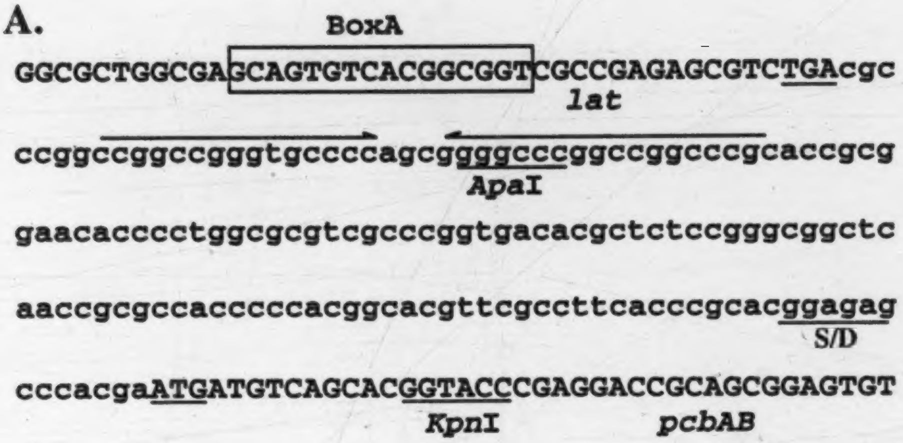
Since the predicted size of the *pcbAB* gene was approximately 11 kb a polycistronic transcript containing *pcbAB* would be very large and therefore it may not be technically possible to determine if a polycistronic transcript including this gene does not exist or if it is just difficult to detect. Kovacevic et al. (123) were able to detect a 10 kb *cefD*- and *cefE*-containing transcript but the *pcbAB*-containing transcript just may be very unstable (175). S1 nuclease protection was employed because the technique is more sensitive than northern analysis and does not rely on the mRNA being of very high quality since the reaction only concentrates on the start of a transcript and not on detecting a number of intact transcripts (97). The S1 probes were designed such that the end-labeled DNA probe fragment will extend upstream beyond the 5' end of the RNA transcript. As well some non-homologous sequence was included at the 3' end of the complementary DNA strand which would make it possible to differentiate between probe reannealing and a probe which is protected by a homologous transcript over its entire length (175).

Using the *lat*up probe which included 227 bp of upstream DNA, S1 nuclease mapping detected a transcription start point (*tsp*) beginning 88 bp upstream of the *lat* start codon (176). The *lat*up promoter fragment gave very strong promoter activities in the promoter probe plasmid pIJ4083 (Table 1), confirming the presence of a promoter within the *lat* upstream region.

Using the *pcbAB*up probe which spanned the *lat-pcbAB* intergenic region and extended 192 bp into the 3' end of the *lat* gene, the *lat-pcbAB* intergenic region was

Figure 4. Organization of the *lat-pcbAB* intergenic region. **A.** Sequence of the *lat-pcbAB* intergenic region contains a predicted stem-loop structure (234) which was overlapped with half arrows. The S/D sequence, the *lat* stop codon, the *pcbAB* start codon and the important restriction sites are underlined. The potential BoxA anti-termination region was enclosed in a box. **B.** The stem-loop structure represents the predicted secondary structure that would occur in the transcript. The *lat* stop codon and the *Apa*I restriction site are underlined.

A.


BoxA
GGCGCTGGCGAGCAGTGTCACGGCGGTCGCCGAGAGCGTCTGAcgc
lat
ccggccggccgggtgccccagcggggcccggccggcccgcaccgcg
ApaI
gaacacccctggcgcgtcgcccggtgacacgctctccgggcggctc
aaccgcgccacccccacggcacgttcgccttcacccgcacggagag
S/D
cccacgaATGATGTCAGCACGGTACCCGAGGACCGCAGCGGAGTGT
KpnI
pcbAB

B.

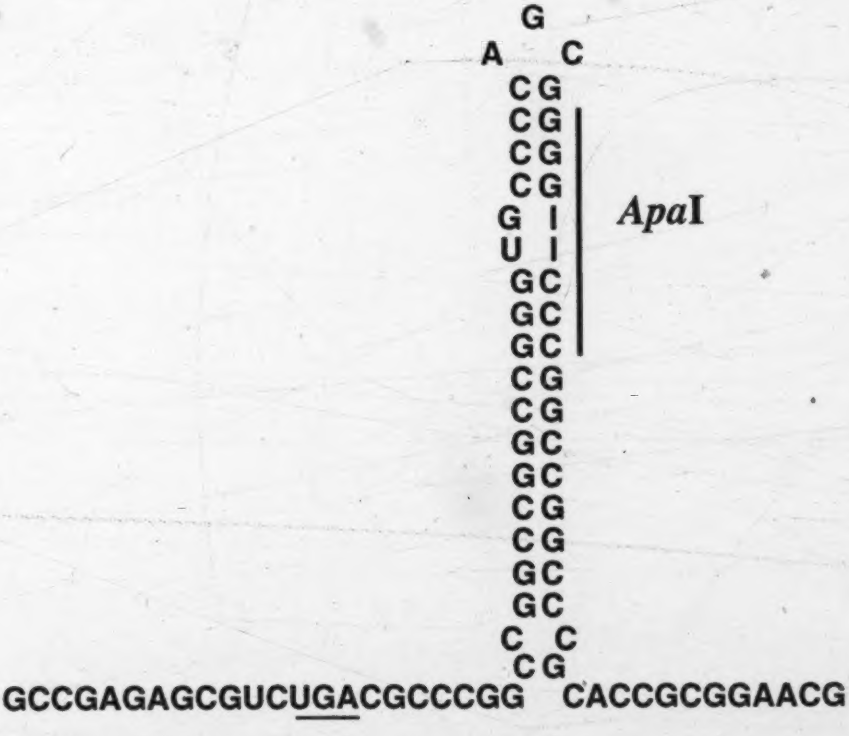

G
A C
C G
C G
C G
C G
G I
U I
G C
G C
G C
C G
C G
G C
G C
C G
C G
G C
G C
C C
C G
GCCGAGAGCGUCUGACGCCCGG CACCGCGGAACG
ApaI

Figure 5. Organization of the *pcbAB-pcbC* intergenic region. **A.** Sequence of the *pcbAB-pcbC* intergenic region contains a region of dyad symmetry which was overlapped with half arrows and would encode an mRNA sequence capable of forming a stem-loop structure (63, 109). The S/D sequence, the *pcbAB* stop codon, the *pcbC* start codon, the predicted -10 and -35 promoter regions are underlined and the transcription start point (*tsp*) determined by high resolution S1 nuclease transcript analysis and primer extension (176, 178) is represented by the arrow. **B.** The stem-loop structure represents the predicted secondary structure that would occur in the transcript. The S/D sequence, the *pcbAB* stop codon and the *pcbC* start codon are underlined.

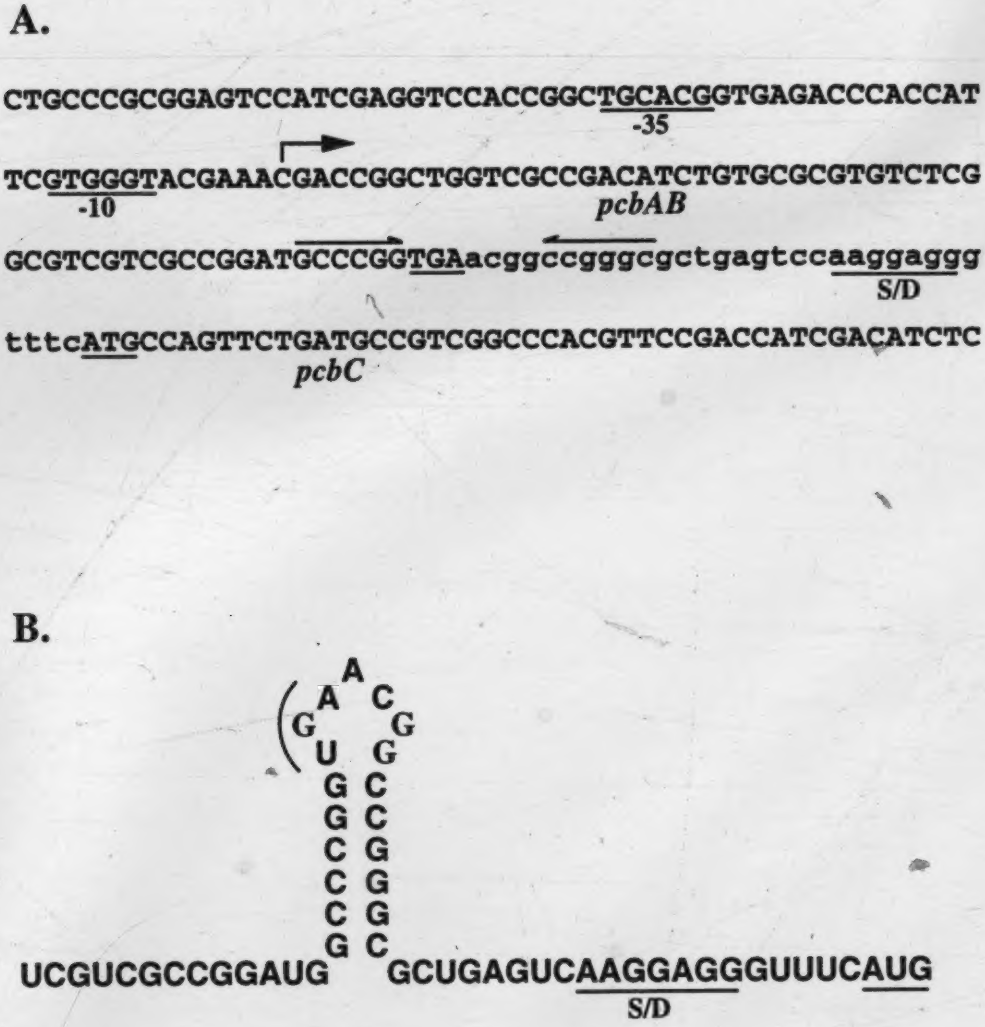

A.
CTGCCCGCGGAGTCCATCGAGGTCCACCGGCTGCACGGTGAGACCCACCAT
-35
TCGTGGGTACGAAACGACCGGCTGGTCGCCGACATCTGTGCGCGTGTCTCG
-10
pcbAB
GCGTCGTCGCCGGATGCCCGGTGAacggccgggcgctgagtccaaggaggg
S/D
tttcATGCCAGTTCTGATGCCGTCGGCCCACGTTCCGACCATCGACATCTC
pcbC
B.
A
A C
G G
U G
G C
G C
C G
C G
C G
G C
UCGUCGCCGGAUG
GCUGAGUCAAGGAGGGUUUCAUG
S/D

Table 1. Promoter probe analysis in *S. lividans* and *S. clavuligerus*[a]

	Catechol dioxygenase specific activity[b]		
plasmid	*S. lividans*	*S. clavuligerus*	*S. clavuligerus* / *S. lividans* [c]
pIJ4083	2	2	1
pIJ4083/*ermE**	295	764	2.6
pIJ4083/*lat*up	34	5260	155
pIJ4083/*pcbAB*up	3	5	0.6
pIJ4083/*pcbC*up	29	77	2.7

[a] Compilation of Table I and Table II in Petrich (175).

[b] C23O specific activity values are given as EU x 10^{-3}/mg protein.

[c] Values are calculated as a ratio of the specific activity obtained for each promoter probe plasmid in *S. clavuligerus* and *S. lividans*.

analyzed by S1 nuclease transcript analysis (176). In all reactions the probes were fully protected over the homologous region suggesting that *lat* and *pcbAB* are present on one contiguous transcript, which had been suggested by the northern analysis. The absence of a *tsp* within this fragment was confirmed by the lack of apparent promoter activity of the *pcbAB*up fragment during promoter probe analysis. Attempts to detect a transcriptional stop point between *lat* and *pcbAB* by S1 nuclease transcript analysis suggested that the stem-loop structure did not create a monocistronic *lat* transcript (175).

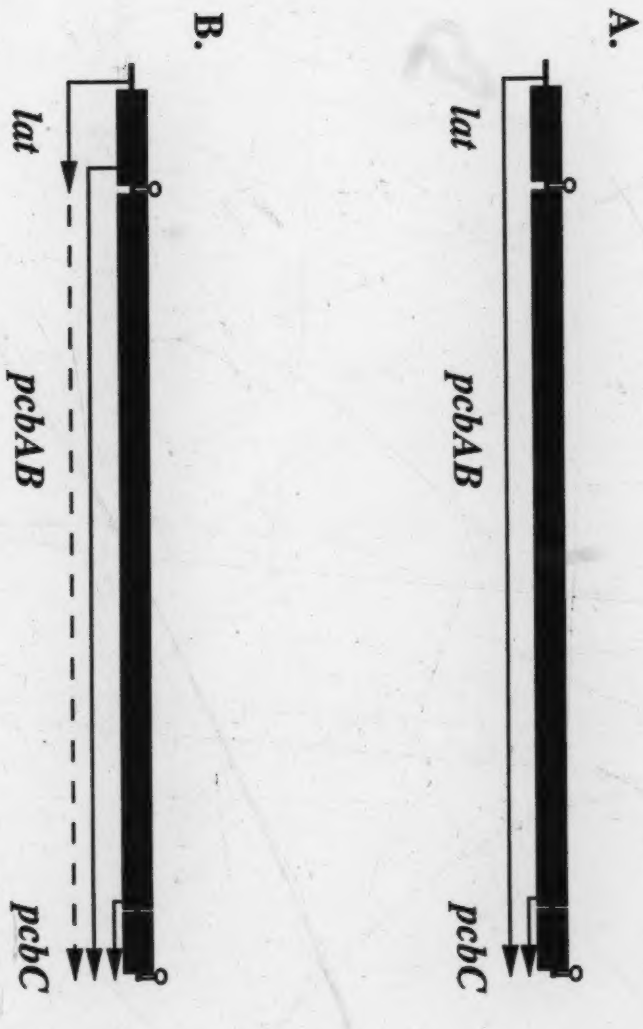
The *pcbC*up probe spanning the *pcbAB-pcbC* intergenic region and extending 307 bp into the *pcbAB* gene 3' end was used for S1 nuclease transcript analysis of the *pcbAB-pcbC* intergenic region (176). A *tsp* was detected 91 bp upstream of the *pcbC* start codon within the *pcbAB* coding region but it was also demonstrated that *pcbC* was part of the *lat*/*pcbAB* polycistronic transcript as an equal percentage of the probe was fully protected. The presence of the *tsp* in the *pcbC*up fragment was confirmed by promoter probe analysis and primer extension analysis (178).

Therefore, S1 nuclease transcript analysis demonstrated transcription originating at the *lat* promoter proceeding into the *pcbAB* gene and then into the *pcbC* gene (176). The co-transcription of these three early genes (Fig. 6A) would provide a means to coordinate control over the onset of antibiotic production (175). As well, a promoter within the *pcbAB* coding region was shown to be responsible for the production of a monocistronic *pcbC* transcript (178).

The NP1 mutant of *S. clavuligerus* generated by N-methyl-N'-nitro-N-nitrosoguanidine (NTG) mutagenesis and characterized by Piret and coworkers (56, 181, 234) produced only 1% of wildtype cephamycin levels. The mutant had normal levels of IPNE and DAOCS but ACVS and IPNS activities had decreased to 17% of wildtype and the LAT enzyme activity had decreased to 4% of wildtype levels (234). The exogenous addition of αAA to cultures increased cephamycin production to 17% of wildtype levels and suggested that the NP1 mutant was a *lat* mutant.

Complementation of the NP1 mutant with a high copy number plasmid containing 20% of the *pcbAB* gene from the 5' end and the *lat* gene minus the *lat* promoter sequences was capable of increasing the cephamycin production from 1% to 67% of wildtype levels (234). Addition of this plasmid increased LAT activity to 72% of wildtype levels but surprisingly this plasmid also caused ACVS and IPNS activity levels to increase to 39% and 35%, respectively. A range of production levels occurred from transformants generated in a single experiment but the majority of these transformants did have increased levels of cephamycin production. Retransformation of plasmid DNA isolated from a low producing strain again yielded a range of production levels and curing the

Figure 6. Transcriptional organization of the *lat*, *pcbAB* and the *pcbC* genes. Coding regions of the genes are indicated by closed boxes and thin lines represent the intergenic regions. Solid arrows represent transcription initiating at the respective promoter and the dashed arrow represents conditional transcription termination. **A.** The co-transcription model proposed by Petrich et al. (176) suggests that the *lat* promoter is responsible for the transcription of the *lat*, *pcbAB* and *pcbC* genes, creating a tricistronic message of approximately 14 kb. As well, the *pcbC* gene was shown to have its own promoter located within the *pcbAB* coding region creating a 1.2 kb *pcbC* monocistronic message (178). **B.** The interdependence model proposed by Yu et al. (234) predicts that the *lat*, *pcbAB* and *pcbC* genes are transcribed as three transcripts originating from the *lat*, *pcbAB* and *pcbC* promoters.


A.
lat
pcbAB
pcbC
B.
lat
pcbAB
pcbC

plasmid from the high producing strain reduced the production back to mutant levels.

From these results it was concluded that an "interdependence" existed between the first three genes for their transcription and that the ACVS and IPNS enzyme activities were dependent upon the production of the LAT protein (56, 234). Their model (Fig. 6B) suggested that αAA would stimulate *pcbAB* transcription to produce the ACVS protein and ACV would in turn stimulate transcription from the *pcbC* promoter to produce IPNS protein.

To determine if the *lat* and *pcbAB* genes were co-transcribed or expressed separately it was necessary to determine whether there was a promoter that could drive *pcbAB* transcription. Yu et al. (234) found that a DNA fragment from the 3' end of the *lat* coding region gave very strong levels of promoter activity in the promoter probe vector pIJ487 and it was suggested that a promoter within this DNA fragment region was responsible for *pcbAB* transcription. Attempts by Petrich et al. (176) to find a *pcbAB* promoter concentrated on the *lat-pcbAB* intergenic region which is downstream of the DNA fragment characterized by Yu et al. (234). Analysis of the DNA fragment containing the putative *pcbAB* promoter showed two possible promoter regions containing sequences very similar to the consensus sequences for *Streptomyces* vegetative promoters. As well, analysis of the *lat-pcbAB* intergenic region revealed a GC-rich stem-loop structure that starts 4 bp downstream of *lat* stop codon.

Attempts were made to demonstrate that the co-transcription model proposed by Petrich et al. (176) rather than the "interdependence" model was the cause for the pleiotropic deficiency of the ACVS and IPNS proteins in the NP1 mutant. The putative *pcbAB* promoter fragment from within the *lat* coding region which gave the strong promoter activities in pIJ487 was also functional in the promoter probe plasmid pIJ4083 but both high resolution and low resolution S1 nuclease mapping of this region failed to detect a *tsp* (169). This suggested that the putative *pcbAB* promoter may not be physiologically active and the strong promoter activity was an artifact of the putative promoter region being on the high copy number plasmid. The *lat-pcbAB* intergenic region contained a putative transcription terminator but a monocistronic *lat* message could not be detected by northern analysis and S1 nuclease mapping (175, 176).

Jensen and coworkers (111) created a *lat* mutant by a gene replacement experiment which would provide further clarification of the validity of the interdependence versus the co-transcription models. The *lat*::*apr* mutant strain resulted from the replacement of the wildtype *lat* gene with a mutant *lat* gene containing the apramycin resistance (*apr*) marker within the *lat* coding region. Since, the *apr* marker was inserted in a direction opposite to the transcription of the *lat* gene and it was expected that the marker would interrupt transcription from the *lat* promoter but not the putative *pcbAB* promoter which was located downstream of the *apr* marker. The *lat*::*apr* mutant did not produce cephamycin but ACVS and IPNS activity were still detected but at reduced levels and the addition of exogenous αAA allowed low amounts of cephamycin to be produced. These data provided support for the interdependence model, however S1 nuclease mapping of RNA isolated from the *lat*::*apr* mutant showed no evidence of a transcript arising from the putative *pcbAB* promoter (169). Furthermore, it could not be determined whether transcription was originating at the *lat* promoter or from a promoter located within the *apr* marker fragment.

Use of the *lat*::*apr* mutant did not demonstrate whether the regulation of *lat*, *pcbAB* and *pcbC* transcription is occurring via the interdependence or co-transcription models. It was felt that additional experiments which could completely eliminate the possibility of transcription from the *lat* promoter were required before the matter could be resolved.

I.6 Regulation of antibiotic production in *Streptomyces*

Studies into the regulation of antibiotic production and morphological differentiation have primarily been conducted in *Streptomyces coelicolor* A3(2) because the organism produces at least four chemically distinct antibiotics, actinorhodin (ACT), undecylprodigiosin (RED), methylenomycin and a calcium-dependent antibiotic. As well, the physiology and developmental cycle are well understood and *S. coelicolor* is the best characterized of the *Streptomyces* spp. in genetic terms (23). It has been suggested that in *S. coelicolor* the regulation of antibiotic production and sporulation is organized into a regulation hierarchy initiating at: 1) regulators common to both sporulation and antibiotic production, such as the *bld* genes, 2) pleiotropic regulators affecting the biosynthesis of two or more antibiotics but not affecting sporulation, such as the *afsR* gene and 3) pathway-specific activators which control the biosynthetic pathway of an individual antibiotic and are located within the cluster of genes, like the *act*II-ORF4, *redD* and *dnrI* genes (140). However, it is more likely that antibiotic production and sporulation are regulated by a regulatory network rather than a regulatory cascade consisting of a number of genes which act together in deciding the appropriateness of commitment to antibiotic production and/or sporulation (44). Action of the regulatory network is carried out via the pathway-specific activators which regulate expression of the biosynthetic genes (74).

Most antibiotics are produced only in the stationary phase of liquid grown cultures or during the onset of morphological differentiation in surface grown cultures which is consistent with a role for growth rate or the cessation of growth in determining the onset of secondary metabolism (44). Mutants defective in ppGpp synthesis were also impaired in antibiotic production which led Ochi (162) to suggest that ppGpp may trigger the onset of antibiotic biosynthesis. Mutants of the *relA* gene which encodes the ppGpp synthetase gene were made in *S. coelicolor* and were shown to be unable to make ppGpp (40, 41). Characterization of the *relA* mutant suggested that *relA* regulation of antibiotic production occurred in a medium dependent manner as cultures grown under nitrogen-limiting conditions failed to make ACT or RED but growth under phosphate-limiting conditions allowed ACT and RED production (40). Under nitrogen-limiting conditions transcription of *redD* was not detected and transcription of *act*II-ORF4 was reduced to a level which may lead to production of an insufficient amount of activator. Nutritional shiftdown during exponential growth initiated by amino acid starvation, a condition which would lead to increased ppGpp synthesis, caused a rapid stimulation of *redD* and *act*II-ORF4 transcription in the wildtype strain but this response was not detected in the *relA* mutant.

Regulated gene expression may reflect changes in the availability of sigma factors. In *E. coli* sets of genes which are only expressed in the stationary phase are dependent upon the *rpoS* encoded sigma factor for expression (93). *Streptomyces* contain multiple sigma factors which confer to the RNA polymerase core enzyme the ability to recognize different subsets of promoter sequences (38). Many genes are expressed from more than one promoter and this sometimes is driven by more than one form of RNA polymerase.

Mutation of the *hrdD* gene had no effect on RED or ACT production even though Fujii et al. (81) determined that RNA polymerase containing σ^{hrdD} directed *in vitro* transcription from the *redD* or *act*II-ORF4 promoter. It was initially thought that these pathway-specific activators are transcribed by a holoenzyme that was different from the major form σ^{hrdB} but the mutation of the *hrdD* gene suggested that the genes are transcribed by another form of the holoenzyme. When three of the minor sigma factor genes similar to the σ^{70} subunit of *E. coli* (*hrdA*, *hrdC* and *hrdD*) were inactivated the resulting triple mutant had no obvious change in the ability to make RED or ACT (37). Mutation of *hrdB*, the major sigma factor gene, was not possible suggesting that the *hrdB* gene was essential (36). Therefore, neither the *red* nor the *act* genes are dependent upon the minor *hrd* genes for transcription and since the -10 and -35 sequences of the *redD* and *act*II-ORF4 genes are similar to the proposed consensus sequence of the major form σ^{hrdB}, transcription may be dependent upon the σ^{hrdB} sigma factor (81). It was suggested that *in vitro* transcription of the *redD* and *act*II-ORF4 promoters may require the addition of a transcriptional activator in addition to the σ^{hrdB} containing RNA polymerase.

Antibiotic production in *Streptomyces* generally coincides with the onset of morphological differentiation and the ability to isolate mutants that are defective in both processes suggest that common elements of genetic control must exist (23). Many different classes of the *bld* mutants exist, most of which are deficient in the ability to sporulate and they also are defective in the biosynthesis of all four antibiotics produced by *S. coelicolor* (42). This class of mutants are referred to *bld*, as the loss of aerial-mycelium formation causes these mutants to appear "bald". The *bldA* mutant is the most extensively characterized of the *bld* mutants and it has been determined that the *bldA* gene is in fact a leucyl tRNA that recognizes the rare UUA codon (132).

It was demonstrated that the mutational loss of the tRNA had no apparent effects on vegetative growth as *bldA* mutants grow rapidly on minimal or rich media but a functional *bldA* gene was essential for antibiotic production and morphological differentiation (137). Analysis of 100 *Streptomyces* genes demonstrated that the TTA codon was present in very few genes and the majority of these genes were involved in aerial mycelial formation, antibiotic resistance or pathway-specific regulation and that these genes are all expressed late in growth (135). The TTA codons were rarely present in biosynthetic genes whereas the other rare leucine codons CTA and TTT are randomly distributed throughout the genes. The majority of the genes which contain a TTA codon have a single TTA codon located in the first 50 amino acids of the protein.

S1 nuclease transcript analysis determined that there were low levels of the mature form of the tRNA in young cultures and this mature tRNA becomes much more abundant in older cultures (137). Northern analysis confirmed the temporal pattern of appearance of the *bldA* tRNA. In contrast, a lysine tRNA which recognized the abundant AAG codon was equally abundant throughout growth. The *carB* gene encodes the carbomycin resistance gene and contains two TTA codons and this prevented a *bldA* mutant strain from acquiring lincomycin resistance when the *carB* gene was introduced (136). Alteration of the two TTA codons in the *carB* resistance gene to CTC codons allowed the *bldA* mutant to become lincomycin resistant. Therefore the absence of the *bldA* tRNA prevents translation of the wildtype *carB* transcript but has no effect on the non-UUA containing mutant *carB* transcript.

The absence of ACT and RED antibiotic production in *bldA* mutants coincided with the transcriptional silence of *act* and *red* genes, respectively (32, 89). The *act*II-ORF4 gene encoding the pathway-specific activator of ACT production contains a TTA codon and therefore translation of the *act*II-ORF4 transcript and biosynthesis of ACT occur in a *bldA*-dependent manner (74). The pathway-specific regulator of RED production, RedD, does not contain a TTA codon and initially it was not known how *bldA* regulated RED production. However, White et al. (225) have recently determined that the *red*Z gene contains a TTA codon and the *red*Z gene was shown to encode a positive regulator required for the transcription of the *redD* pathway-specific activator gene.

bldA regulation however, can be conditional as moderate sporulation of a *bldA* mutant can occur during growth on a poor carbon source like mannitol, but antibiotic production remains *bldA*-dependent under these conditions (148). *bldA* mutants can make RED when grown on low phosphate media but production of other antibiotics and sporulation remained *bldA*-dependent (89). The *hyg* gene encodes the hygromycin phosphotransferase resistance gene and contains a TTA at the seventh codon. When this gene was introduced into a *bldA* mutant at low copy number and the mutant was grown on minimal medium containing mannitol, as a carbon source, resistance to hygromycin was much less than when the gene was introduced into a *bldA*$^+$ strain (136). However, if the *hyg* gene was introduced on a high copy number plasmid, or if the *bldA* strain containing the low copy number plasmid was grown on rich R2YE medium the levels of hygromycin resistance were equal to those observed in the wildtype strain. However, in contrast to the high levels of resistance observed, only low levels of *hyg* gene activity were detected in the *bldA* mutant. This suggested that mistranslation of the UUA codon of the *hyg* transcript occurred allowing the production of low levels of the *hyg* gene product which could support high levels of resistance.

The role of the *bld* gene product regulatory cascade may be to couple morphological differentiation and antibiotic production to the nutritional state of the substrate mycelium (161). The *bld* mutants show a marked defect in catabolite repression of metabolic operons by rich carbon sources as mutants display deregulated expression from the glucose-sensitive galactose-dependent *galP1* promoter (182). Complementation of a *bldA* or *bldB* mutant restored the regulation of expression from the *galP1* promoter, suggesting that the *bld* mutants have a global defect in the regulation of carbon utilization and that the inability of the *bld* mutants to initiate morphogenesis is a secondary consequence of their inability to sense and/or signal starvation. Nodwell and Losick (161) recently have discovered an extracellular signaling molecule referred to a "signal 1" which requires the function of the *bldK* gene product (oligopeptide permease) to be imported into the cell to trigger the next step in the regulatory cascade.

The pathway-specific regulators are often first discovered as a fragment of DNA which stimulates antibiotic overproduction from the wildtype strain. DNA sequencing of the stimulating fragment leads to the discovery of a regulator linked with the biosynthetic genes within the antibiotic biosynthetic cluster and these regulators are able to complement mutants that can not co-synthesize with other non-producing mutants. Regulators have yet to be found for the tetracenomycin, erythromycin or avermectin biosynthesis suggesting that the pathway-specific regulators are not always linked to the biosynthetic clusters (44).

Rather than survey all of the pathway-specific activators which have been discovered in *Streptomyces*, I will concentrate on the *redD* and *act*II-ORF4 genes from *S. coelicolor* as well as the *dnrI* gene from *S. peucetius*, all of which have been very well characterized. The *redD*, *act*II-ORF4 and *dnrI* genes are all homologous to each other as well as to the N-terminus of pleiotropic regulator *afsR* (206). These genes also show similarity to the *ccaR* gene which will be the focus of many of the subsequent studies.

The *act*II-ORF4 gene was first discovered because of its ability to stimulate the overproduction of ACT in *S. coelicolor* and also *S. lividans* which contains both the ACT and RED biosynthetic clusters but in which they are very poorly expressed (74). Mutants created by disruption of the *act*II-ORF4 gene were unable to produce ACT and characterization of RNA isolated from *act*II mutants determined that transcripts corresponding to some of the ACT biosynthetic genes were strongly reduced or absent. Gramajo et al. (86) demonstrated that in wildtype *S. coelicolor* the *act*II-ORF4 transcript accumulated before the *act*III and *act*VI-ORF1 transcripts appeared which would be necessary if ActII-ORF4 is the regulator of their transcription.

The *act*II-ORF4 gene contains a TTA at the fifth codon which would make translation of the *act*II-ORF4 gene product dependent directly on the *bldA* gene and would explain why *bldA* mutants are unable to make ACT (74). *bldA* mutants were demonstrated to produce *act*II-ORF4 transcript but were unable to translate this transcript to allow ACT production (225). Alteration of the TTA codon to a TTG allowed a *bldA* mutant to make high levels of ACT and confirms that the *act* biosynthetic genes do not contain TTA codons (74).

The *redD* gene was first discovered because of its ability to complement and return RED production to blocked mutants which were unable to co-synthesize with other *red* mutants (73). Characterization of the *redD42* mutant determined that the mutant lacked certain biosynthetic enzymes (73) and transcripts for the *redE* and *redBF* genes were reduced when compared to wildtype (155). It was also possible to demonstrate that the appearance of the *redD* transcript preceded that of the transcripts for the *redE* and *redBF* genes in wildtype *S. coelicolor*, which would be necessary if RedD is the regulator of their transcription.

Like *act*II-ORF4 the *redD* gene was able to stimulate overproduction of RED in wildtype *S. coelicolor* or in the *redD42* mutant (155) and the presence of the *redD* gene on a high copy number plasmid resulted in the overproduction of RED during exponential growth (207). Both *redD* and the biosynthetic gene *redX* transcripts were detected earlier and at higher levels when multiple copies of *redD* were added. These results suggest that RED production is limited by the availability of enough RedD activator.

Unlike *act*II-ORF4 the *redD* gene does not contain a TTA codon and it was not known how RED production was regulated by *bldA* (207). *bldA* mutants were demonstrated to be unable to produce the *redD* transcript which suggested that a gene required for *redD* expression must contain a TTA codon (225). The *redZ* gene discovered within the RED biosynthetic cluster displayed homology to response regulators from two-component regulator families and also contained a TTA codon. Expression of a *redZ* transcript was demonstrated in a *bldA* mutant but in the absence of the tRNA it was unable to be translated and initiate RED production.

Mutants of the *redZ* gene were created by gene replacement and these mutants were unable to make RED but were not impaired in their ability to make ACT (225). Analysis of RNA from the *redZ* mutant determined that *redD* transcription was absent which suggested that *redZ* was required for *redD* transcription. However, *redZ* transcription was present in a *redD* mutant and the addition of a high copy number plasmid containing the *redZ* gene was unable to return RED production to a *redD* mutant.

Attempts were made to titrate out the RedZ protein by the addition of the *redD* promoter region on a high copy number plasmid (225). Addition of the *redD* promoter did not result in a decrease in the amount of RED produced and as well the addition of multiple copies of the *redD* gene resulted in increased RED production. Both of these results suggest that the amount of RedZ was not limiting. Analysis of RNA isolated from wildtype *S. coelicolor*, *S. coelicolor* with *redD* on a high copy number plasmid or the *redZ* mutant demonstrated that great increases in the amount of *redZ* transcript occurred when extra copies of a RedZ-dependent promoter were added or when a functional RedZ protein was not produced. Both of these conditions would reduce the autoregulation of the *redZ* promoter and the resulting amount of *redZ* transcript would increase. Since antibodies specific for RedZ had not been generated it was not possible to determine if the multiple copies of the *redD* promoter region actually resulted in the production of greater amounts of RedZ protein to compensate for the increased number of RedZ targets.

As mentioned earlier RED production can be restored to a *bldA* mutant by reducing the concentration of phosphate but the mutants remain deficient in ACT production and morphological differentiation. White and Bibb (225) determined that functional *redD* and *redZ* genes were both required for the restoration of RED production. It is believed that a *bldA* mutant can produce small amounts of the RedZ protein through mistranslation of the UUA codon within the *redZ* transcript.

Stutzman-Engwall et al. (206) found two regulatory regions in the daunorubicin (DNR) cluster of *S. peucetius* which resulted in DNR overproduction when transformed into wildtype *S. peucetius*. Sequence analysis of the first of these regions led to the discovery of the *dnrI* gene, which showed homology to *Streptomyces* pathway-specific regulators like RedD and ActII-ORF4. Mutants created by gene replacement of the *dnrI* gene did not produce DNR or any of the biosynthetic intermediates. These *dnrI* mutants also had increased sensitivity to high levels of DNR and complementation of the *dnrI* mutants resulted in a return of high level DNR resistance.

Transcriptional analysis of the *dnrI* mutants determined that the absence of a functional *dnrI* gene prevented transcription of the putative operons containing the *dnmLM*, *dnrXY*, *dnmZUV*, *dpsABCD* and *dnrDKPQS* biosynthetic genes which are involved in daunosamine biosynthesis, polyketide biosynthesis or modifications to the final antibiotic (140). The absence of transcription from these genes would obviously explain the absence of DNR production. As well, the *dnrI* mutants lacked the transcripts for the *drrAB* and *drrC* genes required for resistance to DNR.

The lack of DNR intermediates and transcription of the biosynthetic and resistance genes suggests that the *dnrI* gene coordinates control over antibiotic production and resistance (209). Attempts to overexpress DnrI in wildtype *S. peucetius* were unsuccessful but the overproduction was successful in the *dpsB* mutant strain which would not accumulate any toxic DNR intermediates. DnrI protein purified from *S. peucetius* or *E. coli* was able to bind to the promoter regions of early and late genes in gel-shift experiments. DnaseI footprinting experiments determined that the DnrI protein would bind to an imperfect inverted repeat which contains a 5'-TCGAG-3' consensus sequence located just upstream of the -35 region of the promoter. This consensus sequence was found upstream in 4 out of 10 other promoter regions in the DNR cluster.

The *dnrI* gene was able to behave like the homologous *act*II-ORF4 gene from *S. coelicolor* as the *dnrI* gene could complement an *act*II-ORF4 mutants to return actinorhodin production (206). Wietzorrek and Bibb (228) were able to find promoters within the *act* biosynthetic cluster which also contained the 5'-TCGAGCG/C-3' or 5'-TCGAGCG-3' consensus sequence repeats separated by 11 or 22 bp of adjacent sequence, which would correspond to one or two complete turns of the DNA helix and therefore would be present on the same face of the DNA.

Sequence analysis of the second regulatory region in the DNR cluster, which had given overproduction of DNR by the wildtype strain, led to the discovery of the *dnrN* gene which showed homology to the response regulators from two-component regulator families (163). A gene homologous to the histidine kinase that would act as the sensor and phosphorylate DnrN in response to an appropriate stimuli has not been found within the DNR biosynthetic cluster (83). Mutants created by gene replacement of the *dnrN* gene were phenotypically similar to the *dnrI* mutants as they did not produce DNR or any of the biosynthetic intermediates as well as having increased sensitivity to high levels of DNR (163).

Transcriptional analysis of the *dnrN* mutants determined that the absence of a functional *dnrN* gene prevented transcription of the *drrAB* and *drrC* resistance genes as well as the *dnrI* gene (140). It was suggested that a regulatory cascade existed where DnrI regulates transcription of the resistance and biosynthetic genes and the DnrN protein regulates *dnrI* expression. It was possible to complement the *dnrN* mutant by adding the *dnrN* gene *in trans* but the addition of *dnrN* gene *in trans* was unable to complement a *dnrI* mutant. Tang et al. (209) demonstrated that altering the promoter of the *dnrI* gene allowed expression in a *dnrN* mutant background.

The DnrN protein also displays some temporal regulation as the DnrN protein peaks at the onset of DNR production and then begins to decrease as DNR accumulates (83). The DnrN protein could be demonstrated to bind to only the *dnrI* promoter region specifically by a gel-shift assay. DNaseI footprinting analysis demonstrated two regions of specific binding within the *dnrI* promoter and both contain consecutive overlapping triplets which have also been shown to be the preferential binding sites for DNR suggesting that high levels of DNR may downregulate transcription of the *dnrI* gene. This would be the first example of the product of the biosynthetic pathway playing a regulatory role in its own production.

Attempts were made to titrate out the DnrN protein by the addition of the *dnrI* promoter region on a high copy number plasmid (83). Addition of the *dnrI* or *dnrN* promoters did not result in a decrease in the amount of DNR produced and suggested that the regulatory proteins must not bind tightly enough or the plasmid copy number was not high enough.

A number of similarities exist among the various pathway-specific regulators as the stationary phase production of the antibiotics initiates with the transcription of the biosynthetic genes which were silent until transcription of the pathway-specific activator occurred. Overexpression of these genes on high copy number plasmids results in the amplification of the amount of antibiotic produced and causes antibiotic production to initiate earlier than normal. Mutants of the pathway-specific activators are unable to produce antibiotic under any conditions and analysis of RNA isolated from these mutants demonstrates an absence or reduction in the transcripts of the biosynthetic genes.

The *redD* and *dnrI* genes are unusual in that these pathway-specific regulators are directly regulated at the level of transcription by other pathway-specific regulators located within the biosynthetic cluster. The *redD-redZ* relationship explains how RED production is regulated in a *bldA* mutant when the *redD* gene does not contain a TTA codon. It had always been assumed that these homologous pathway-specific activators would directly bind to the regulated promoters but the absence of a helix-turn-helix motif within primary amino acid sequence left the mechanism unclear. Work on the regulation of DNR biosynthesis provided the first demonstration of DNA-binding and therefore direct regulation by the pathway-specific regulators at promoters that were DnrI- or DnrN-dependent (83, 209). This along with the work demonstrating that these homologous proteins may contain and OmpR-like DNA-binding fold which would allow these proteins to bind to DNA (228) provides an explanation for how these pathway-specific activators positively regulate transcription of the cognate biosynthetic structural genes.

There are also a number of pleiotropic regulators [*absA* (2), *absB* (1), *aba* (75), *afsQ*$_1$/*Q*$_2$ (101) and *afsR* (98)] which are not pathway-specific because they affect the production of more than one antibiotic, but they are not global regulators because the mutations do not affect sporulation. The *afsR* gene was first discovered because it led to high level production of both ACT and RED when added to *S. coelicolor* or *S. lividans* (98). The *afsR* gene encodes a very large protein, which contains two putative helix-turn-helix motifs, two ATP consensus binding sequences (99) and the N-terminal end 1/3 of the gene is homologous to the pathway-specific activators *redD*, *act*II-ORF4 and *dnrI* (77). Addition of multiple copies of the *afsR* gene to a *redD* mutant resulted in overproduction of ACT and similarly, addition of the *afsR* gene to an *act*II-ORF4 mutant resulted in overproduction of RED. The AfsR protein however could not restore any lost activity by replacing the missing pathway-specific activator as the addition of multiple copies of the *afsR* gene to a *redD* mutant did not return RED production to any level. The addition of multiple copies of *afsR* did however result in increased levels of *act*II-ORF4 and *redD* transcripts which was especially profound during the late exponential phase when these transcripts were normally barely detected.

The AfsR protein is phosphorylated by a membrane-associated serine-threonine protein kinase called AfsK, suggesting an important role for signal transduction via phosphorylation in determining the onset of antibiotic production (147). Disruption of the *afsK* gene resulted in significant but not complete loss of ACT production. Disruption of the *afsR* gene resulted in a similar phenotype on minimal medium where there was a significant decrease in both ACT and RED production (77). Levels of the *act*II-ORF4 and *redD* transcripts were not decreased by the mutation of *afsR* and the addition of multiple copies of the *act*II-ORF4 or *redD* genes resulted in either ACT or RED overproduction, respectively. On rich media like, R2 or R2YE, production of ACT and RED were very similar to wildtype. Alteration of the phosphate concentration in R2 media returned a decrease in RED production suggesting that *afsR* may be required only under certain conditions.

A detailed molecular model for how antibiotic production and sporulation is regulated in various *Streptomyces* spp. is lacking but it appears that there are many different regulatory effectors within the overall network whose actions overlap but are focused upon the temporal regulation of the pathway-specific activators. It is unlikely that all of these regulators are functioning in all organisms and/or under all conditions and *S. coelicolor* may be an example of regulation to its fullest degree because of the multiple antibiotics produced by this organism. Information as to how the regulation of antibiotic biosynthesis occurs would provide insight towards means with which to manipulate industrially important organisms to improve antibiotic production. The study of regulation of antibiotic biosynthesis in these organisms may be best served by initially finding the pathway-specific activators and then working backwards to try and determine factors which are responsible for the temporal regulation of expression observed for the pathway-specific activators.

I.7 Regulatory genes within the *S. clavuligerus* cephamycin cluster

Promoter probe analysis of the *lat*up promoter fragment showed very strong promoter activity in *S. clavuligerus* but the activity was 155-fold weaker in *S. lividans* (see Table 1; 175). The species-specific activation of the *lat*up promoter fragment in *S. clavuligerus* suggested the presence of a positive regulatory factor acting specifically upon the *lat* promoter.

Barton and coworkers (20, 221) reported a class of clavulanic acid non-producing mutants generated during strain improvement which could not co-synthesize with other non-producing mutants and were also unable to produce cephamycin. The biosynthesis of both antibiotics could be recovered by complementation. Sequencing of this complementing fragment led to the discovery of the first transcriptional activator, called *dclX* for d̲ecreased c̲l̲avulanic acid, responsible for the regulation of β-lactam biosynthesis. The DclX protein sequence showed homology to many pathway-specific transcriptional activators from different *Streptomyces*; such as RedD, ActII-ORF4, DnrI and the N-terminus of the AfsR protein.

This transcriptional activator was later described by Perez-Llarena et al. (174) and called CcaR for c̲ephamycin and c̲lavulanic a̲cid biosynthesis r̲egulator. The *ccaR* gene was shown to be essential for the biosynthesis of cephamycin and clavulanic acid as mutants of *ccaR* had lost the ability to produce cephamycin or clavulanic acid. When the CcaR primary amino acid sequence was aligned with other homologous activators (RedD, ActII-ORF4, AfsR and DnrI) many regions of homology were present throughout the primary sequence. The CcaR protein also contains the rare leucine encoding TTA codon often found in these pathway-specific activators. We prefer the CcaR nomenclature over DclX since CcaR designates the regulation of both cephamycin and clavulanic acid biosynthesis and this nomenclature will be used throughout the study. The gene will only be referred to a *dclX* when designating the different start codon chosen (221) or when referring to its timing of discovery.

Since the CcaR protein shows homology to other pathway-specific activators it likely functions in a manner similar to these proteins. Attempts to begin understanding the regulation mechanism of β-lactam antibiotic production required that mutants of the *ccaR* gene be generated by gene replacement. The resulting mutants were then characterized for production of the various β-lactam antibiotics produced by *S. clavuligerus* and the enzymes required for cephamycin biosynthesis. *S. clavuligerus* mutants generated by random mutagenesis which lacked the ability to produce the CcaR protein, cephamycin, as well as many of the cephamycin biosynthetic enzymes, were characterized as potential regulatory mutants.

II. Materials and Methods

II.1 Materials

Restriction enzymes were obtained from New England Biolabs Ltd. (Mississauga, ON), Promega Corp. (Madison, WI) or Boehringer Mannheim Biochemicals (Laval, QC) and used according to manufacturers' recommendations. S1 nuclease, exonuclease III (exoIII), T4 DNA ligase, T4 DNA polymerase, calf intestinal alkaline phosphatase, Klenow fragment of *E. coli* DNA polymerase, hygromycin (hyg), random hexanucleotide primers, deoxynucleoside triphosphates (dNTPs) as lithium salts, glycogen, lambda DNA, Taq DNA polymerase and ultrapure Tris(hydroxylmethyl)-aminomethane (Tris) buffer were also obtained from Boehringer Mannheim Biochemicals. Amylose resin, Factor Xa protease, and Vent DNA polymerase were obtained from New England Biolabs Ltd. The Erase-A-Base exoIII deletion kit and Wizard miniprep columns were purchased from Promega Corp. Electroporation cuvettes were purchased from BTX Inc. (San Diego, CA). GlassMax DNA isolation system, high range prestained protein molecular weight standards and electrophoresis grade sodium dodecyl sulfate (SDS) were obtained from Gibco BRL (Burlington, ON).

The PVDF membrane (Immobilon-P) used for western transfers was obtained from Millipore Corp. (Bedford, MA) and the nylon membrane (Hybond-N) used in Southern transfers was purchased from Amersham Life Science Inc. (Oakville, ON). Southern blots were exposed to Kodak X-OMAT AR X-ray film (Eastman Kodak Co., Rochester, NY) while developed western blots were exposed to Fuji RX X-ray film (Fuji Photo Film Co., Tokyo, Japan). Western blots were developed using the ECL western blotting detection reagents and horseradish peroxidase conjugated donkey anti-rabbit Ig secondary antibody both of which were purchased from Amersham Life Science Inc.

Bis-δ-(L-α-aminoadipyl-L-cysteinyl)-D-valine (bis-ACV) was purchased from Peptides International Inc. (Louisville, KY). Chemically synthesized isopenicillin N and penicillin N were obtained from S. Wolfe (Simon Fraser University, Burnaby, BC).

Catechol, dimethylsulfoxide (DMSO), dithiothreitol (DTT), L-α-aminoadipic acid, isopropyl β-D-thiosgalactopyranoside (IPTG), N-tris[hydroxymethyl]methyl-2-aminoethanesulfonic acid (TES) buffer, 3-[N-morpholino]propanesulfonic acid (MOPS) buffer, imidazole, lysozyme, ampicillin (amp), thiostrepton (tsr), kanamycin (kan), and neomycin sulfate (neo) were all purchased from Sigma Chemical Corp. (St. Louis, MO). Apramycin (apr) was purchased as Apralan from PROVEL Division Eli Lilly Canada Inc. (Scarborough, ON). Polyethylene glycol (PEG) 1000 was purchased from Koch-Light

Ltd. (Haverhill, England). Electrophoresis grade glycine was purchased from ICN Biomedicals Inc. (Cleveland, OH). All other chemicals used were of reagent grade.

Trypticase soy broth was purchased from Becton Dickinson and Company (Cockeysville, MD). Bacto-peptone, soluble starch, nutrient broth, yeast extract, malt extract, casamino acids, skim milk, maltose, bacto-penase, ISP medium 3 (ISP3) and bacto-agar were obtained from Difco Laboratories (Detroit MI). Universal bottles were obtained from United Glass (United Kingdom).

Custom oligonucleotide primers were purchased from the Department of Biological Sciences DNA Synthesis Laboratory, University of Alberta, Edmonton, Alberta. Radioactively labeled [α-^{32}P] dATP or dCTP were obtained from Amersham Life Science Inc.

II.2 *Streptomyces* procedures

All manipulations involving *Streptomyces* cultures were done aseptically using sterile solutions, tubes, pipettes, etc., except for the final harvesting of production cultures. All centrifugation procedures were done using a JA-20 rotor in an Beckman centrifuge at 8000 rpm and cooled to 4°C, unless otherwise noted. All microcentrifugation steps were done at room temperature in an Eppendorf 5415 C microcentrifuge (Brinkmann Instruments Inc., Mississauga, ON) at full speed (14000 rpm), unless otherwise noted.

II.2.1 Storage of strains

Wildtype *S. clavuligerus* and chromosomally altered strains (gene replacement mutants or strains containing the integrating vector pSET152) were maintained as glycerol spore stocks. Spores were generated on ISP3, which was supplemented with antibiotic when necessary. Plates were incubated for 12 to 16 days before the spores were harvested in 0.1% (w/v) peptone water. Spore suspensions were sonicated for 5 to 10 minutes in a sonicator bath before passing through a syringe containing sterile non-absorbent cotton. Spore suspensions were then microcentrifuged for 20 seconds, decanted, resuspended in 20% (w/v) glycerol and stored at -70°C.

Cultures of *S. clavuligerus* containing plasmids were stored as glycerol stocks in which equal volumes of culture and 50% (w/v) glycerol were mixed and stored at -70°C. The *S. clavuligerus* NTG mutants were also stored as glycerol stocks because the mutants did not sporulate very well. A list of the bacterial strains used or created in this study is provided in Table 2.

Table 2. Bacterial strains list.

Strains	Description
E. coli	
HB101	Used for general cloning experiments and as a recipient for transformation of shuttle plasmids purified from *Streptomyces*. (From New England Biolabs Ltd., Mississauga, ON)
DH5α	Used for general cloning experiments requiring α-complementation. (From New England Biolabs)
ER1447	The strain is *dam*$^-$ and *dcm*$^-$ and was used to reduce the methylation state of the plasmids bound for *Streptomyces*. (From R. Losick)
ESS	β-lactam supersensitive strain used as an indicator organism for the penicillin/cephalosporin antibiotic bioassays. (From A. Demain)
M. luteus	
ATCC 9341	Indicator organism for the IPNS bioassays. (From ATCC)
S. clavuligerus	
3585	The wildtype strain capable of production of cephamycin C, clavulanic acid and non-clavulanic acid clavams. (From NRRL)
*dclX*68	Cephamycin and clavulanic acid non-producing derivative of wildtype *S. clavuligerus* created by random mutagenesis. Strain supplied only as genomic DNA since it is a derivative of a high-production industrial strain. [From B. Barton, (221)]
NTG 19	Cephamycin non-producing derivative of wildtype *S. clavuligerus* created by random NTG mutagenesis. [Jensen (106), previous study]
NTG 26	Cephamycin non-producing derivative of wildtype *S. clavuligerus* created by random NTG mutagenesis. Also an auxotroph requiring the addition of 0.2% (w/v) casamino acids to SA medium. [Jensen (106), previous study]
NTG 32	Cephamycin non-producing derivative of wildtype *S. clavuligerus* created by random NTG mutagenesis. [Jensen (106), previous study]
NTG 83	Cephamycin non-producing derivative of wildtype *S. clavuligerus* created by random NTG mutagenesis. ([Jensen (106), previous study]
ccaR::apr	*ccaR*$^-$ derivative of wildtype *S. clavuligerus* created by a gene disruption experiment. The mutant *ccaR* gene contains the *apr* cassette inserted into the *Eco*ICRI site. (This study)
Δ*ccaR::tsr*	*ccaR*$^-$ derivative of wildtype *S. clavuligerus* created by a gene replacement experiment. The wildtype *ccaR* gene was deleted and replaced with the *tsr* resistance marker. (This study)

Table 2. (continued)

Strains	Description
S. clavuligerus	
Δ*cob*::*tsr*	*ccaR*⁻, *orf11*⁻ and *blp*⁻ derivative of wildtype *S. clavuligerus* created by a gene replacement experiment. The wildtype genes were deleted and replaced with the *tsr* resistance marker. (This study)
orf11::*apr*	*orf11*⁻ derivative of wildtype *S. clavuligerus* created by a gene disruption experiment. The mutant *orf11* gene contains the *apr* cassette inserted into the *Nco*I site. (This study)
Δ*lat*::*tsr*	*lat*⁻ derivative of wildtype *S. clavuligerus* created by a gene replacement experiment. The mutant *lat* gene contains the *tsr* resistance marker replacing a small internal deletion of the *lat* gene. (This study)
Δ*lat*::*tsr*/term	*lat*⁻ derivative of wildtype *S. clavuligerus* created by a gene replacement experiment. The mutant *lat* gene contains the *tsr* resistance marker and FKMT terminator replacing the *lat* promoter and part of the *lat* gene. (This study)
lat::*apr*	*lat*⁻ derivative of wildtype *S. clavuligerus* created by a gene disruption experiment. The mutant *lat* gene contains the *apr* resistance marker inserted into the unique *Kpn*I site. [Jensen et al. (111), previous study]
Δ*lat*	*lat*⁻ derivative of *lat*::*apr* created by a second gene replacement experiment. The *apr* disrupted *lat* gene was replaced with a mutated *lat* gene containing an internal deletion. (This study)

II.2.2 Introduction of DNA into *Streptomyces*

II.2.2.1 Preparation of protoplasts from *S. clavuligerus* strains

Spores or mycelial stocks were inoculated into 10 ml of trypticase soy broth + 1% (w/v) maltose (TSBM), containing antibiotic(s) when necessary, in a universal bottle with a stainless steel spring. The bottles were incubated for 24 to 36 hours at 28°C on a rotary shaker (280 rpm) to give heavy growth. One milliliter of seed culture was then transferred into a 125 ml flask with a spring containing 10 ml of TSBM and 15 ml of YEME [0.3% (w/v) yeast extract, 0.5% (w/v) peptone, 0.4% (w/v) malt extract, 1% (w/v) maltose and 34% (w/v) sucrose (97)] and the total 25 ml was supplemented with $MgCl_2$ to 5 mM and glycine to 0.5% (w/v). After incubating for 18 to 24 hours the culture was transferred into a sterile universal bottle and centrifuged for 10 minutes at a setting of 2.9 in a PRJ centrifuge before twice washing the mycelia with 10 ml of 10.3% (w/v) sucrose. At this point the mycelial pellet could be frozen at -20°C for later use or resuspended in 4 ml of P buffer [10.3% (w/v) sucrose, 0.057 mM K_2SO_4, 10 mM $MgCl_2$, 0.005% (w/v) KH_2PO_4, 0.368% (w/v) $CaCl_2$, 0.573% (w/v) TES pH 7.2, 0.59 μM $ZnCl_2$, 1.48 μM $FeCl_3$, 0.12 μM $CuCl_2$, 0.10 μM $MnCl_2$, 0.052 μM $Na_2B_4O_7$ and 0.016 μM $(NH_4)_6Mo_7O_{24}$ (97)] plus 2 mg/ml lysozyme and incubated at room temperature for 15 minutes. The mycelial suspension was triturated three times with a 10 ml pipette and incubated for another 5 minutes before 5 ml of P buffer was added. The protoplasts were filtered through a 10 ml syringe containing a plug of sterile non-absorbent cotton and collected in a snap-cap centrifuge tube as an additional 5 ml of P buffer was added to wash the cotton. Protoplasts were centrifuged for 10 minutes at 3000 rpm and washed once with 10 ml of P buffer before resuspending the pellet of protoplasts in 2 ml of P buffer. Protoplasts were dispensed into 1.5 ml centrifuge tubes as 200 to 250 μl aliquots and slowly frozen at -70°C in a container of ice or used for transformations immediately.

II.2.2.2 Transformation of *S. clavuligerus*

Transformation and selection of transformants was carried out as a modification of the methods described in Hopwood et al. (97) and Bailey and Winstanley (13). Frozen protoplasts that were quick thawed or fresh protoplasts were supplemented with 500 μl of P buffer and centrifuged for 7 minutes at 3000 rpm in a microfuge at room temperature. After discarding the supernatant, the protoplasts were resuspended in the remaining drop of P buffer and heat shocked by incubating the tube at 45°C for 5 minutes. Up to 20 μl of plasmid DNA was added to the protoplasts and followed immediately by 100 μl of T buffer (P buffer + 25% (w/v) PEG 1000) and then 1 ml of P buffer before the protoplasts were centrifuged as above. After discarding the supernatant, the pellet of protoplasts was resuspended and 500 µl of P buffer was added before aliquots of the transformation mixture were plated out on dried 20 ml modified R5B plates [10% (w/v) sucrose, 1% (w/v) soluble starch, 0.1% (w/v) casamino acids, 65 mM sodium glutamate, 0.2 mM $MgSO_4$, 25 mM $MgCl_2$, 0.005% (w/v) KH_2PO_4, 0.368% (w/v) $CaCl_2$, 0.573% (w/v) TES pH 7.2, 0.59 µM $ZnCl_2$, 1.48 µM $FeCl_3$, 0.12 µM $CuCl_2$, 0.10 µM $MnCl_2$, 0.052 µM $Na_2B_4O_7$, 0.016 µM $(NH_4)_6Mo_7O_{24}$ and 2.7% (w/v) agar (13)].

The plasmid DNA used for the *S. clavuligerus* transformations was isolated from *E. coli* ER1447 or other strains of *S. clavuligerus* and resuspended in TE (10 mM Tris-HCl pH 8.0 and 1 mm EDTA) without RNaseA. The pSET152-based, pDA2000-based plasmids or smaller shuttle plasmids purified from *E. coli* ER1447 could be transformed directly into *S. clavuligerus*. Larger plasmids, like the gene targeting shuttle vectors, were transformed into *S. clavuligerus* as DNA isolated from other *S. clavuligerus* strains.

Plates were incubated at 28°C for 36 to 48 hours for *S. clavuligerus* transformations before the plates were flooded with 2.5 ml of soft nutrient agar [0.8% (w/v) nutrient broth and 0.7% (w/v) agar (97)] plus antibiotic for a 22.5 ml total volume. Final antibiotic concentrations were 5 µg/ml, 25 µg/ml and 50 µg/ml for tsr, apr and neo, respectively. Hyg was never used to overlay transformation plates but was used at 50 µg/ml or 200 µg/ml when used in liquid or solid media, respectively. In many cases the gene targeting vectors contained more than one selectable marker but the transformation plates were never overlayed with more than one antibiotic. After allowing the overlay to solidify, the plates were returned to the 28°C incubator until transformants appeared, usually 2 to 5 days.

Colonies that arose from transformations were patched out onto MYM agar plates [0.4% (w/v) maltose, 0.4% (w/v) yeast extract, 1.0% (w/v) malt extract and 1.8% (w/v) agar (205)] plus antibiotic as appropriate. The patch plates were then incubated at 28°C for 3 to 6 days and positive transformants were judged as those transformants which grew well on the MYM plates containing antibiotics. Each of the two transformants were inoculated into 6 ml of TSBM plus antibiotic(s) in universal bottles with springs. If the transformants were for *xylE* promoter probe analysis, a minimum of three transformants were screened due to variations in expression, and then the two transformants which gave similar expression levels were chosen for further study. The transformants were grown for up to 48 hours before an aliquot was removed to make a glycerol stock and if required, aliquots were taken to isolate plasmid, streak for spores or to bioassay for antibiotic.

II.2.3 Analysis of DNA from *Streptomyces*

II.2.3.1 Plasmid DNA isolation

Isolation of *Streptomyces* plasmids was accomplished using a modification of the *E. coli* method (28) designed specifically for *Streptomyces* (21). Plasmid DNA was isolated by centrifuging 1.5 ml of culture and washing the mycelia pellet with 200 µl of solution I (25 mM Tris-HCl pH 8.0, 50 mM glucose and 10 mM EDTA) before resuspending in 200 µl of solution I supplemented with 4 mg/ml lysozyme and 10 µg/ml RNaseA. Tubes were incubated at 37ºC for 30 to 60 minutes before 400 µl of solution II (1% SDS, 0.2 N NaOH) was added, mixed by inversion and incubated for 5 minute on ice. Three hundred microlitres of solution III [3 M potassium acetate (KOAc) pH 4.8] was added and mixed by inversion before placing on ice for 5 minutes. The tubes were mixed well by both vortexing and inversion after the addition of 400 µl of phenol/chloroform/isoamyl alcohol (25:24:1) and then microcentrifuged for five minutes before 800 µl of upper phase was transferred into fresh tubes containing 800 µl of isopropanol. After a 10 minute microcentrifugation the pellets were washed with 200 µl of 70% ethanol and vacuum dried.

Shuttle plasmid DNA was transformed into *E. coli* HB101 and the DNA isolated from *E. coli* was analyzed for plasmid rearrangements by restriction enzyme digestions.

II.2.3.2 Genomic DNA isolation

Isolation of *Streptomyces* genomic DNA was accomplished using a modification of the small scale total DNA isolation procedure (97). Genomic DNA was isolated by centrifuging 1.5 ml of culture and washing twice with 500 µl of 10.3% (w/v) sucrose. The washed pellets were resuspended in 500 µl of lysozyme buffer (25 mM Tris-HCl pH 8.0, 0.3 M sucrose, 25 mM EDTA, 2 mg/ml lysozyme and 50 µg/ml RNaseA) and incubated at 37ºC for 30 to 45 minutes. Two hundred and fifty microlitres of 2% (w/v) SDS was added to the samples which were vortexed vigorously for approximately 20 seconds after which 250 µl of phenol/chloroform/isoamyl alcohol was added and the samples were vigorously vortexed for 1 minute. The tubes were microcentrifuged for 5 minutes and the upper phase was re-extracted until the interface was clear. The upper phase was then extracted with an equal volume chloroform/isoamyl alcohol (24:1) and centrifuged. The upper phase was added to a fresh tube and mixed with 0.1 volume of 3 M sodium acetate (NaOAc) pH 5.2 and 2 volumes of 98% ethanol and inverted a number of times. After sitting on ice for 20 minutes the tubes were microcentrifuged for 10 minutes, the supernatants discarded and the pellets washed with 70% ethanol. Pellets were then vacuum dried.

II.2.3.3 Southern blotting

The transfer of DNA from agarose gels to nylon membranes was a modification of the method of Southern (199) as described in Rigaud et al. (187). Restriction enzyme digested genomic DNA was separated on 0.7% or 1.0% agarose gels with TAE running buffer (40 mm Tris-acetate pH 8.0, 1 mM EDTA). After photography, the gel was soaked twice for 10 minutes in 0.25 M HCl, rinsed briefly with water and soaked twice for 15 minutes in denaturing buffer (0.5 M NaOH and 1 M NaCl). The gel was rinsed well with water again and then soaked twice for 30 minutes in neutralization/transfer buffer (1 M NH_4OAc and 0.02 N NaOH). The gel was placed well side down on a piece of filter paper that had been placed on top of the wick soaked in transfer solution. The nylon membrane was laid on top of the gel followed by two pieces of filter paper. The filter paper and the nylon membrane were moistened in transfer buffer before using. A stack of paper towels and a weight were then placed on top of the sandwich and left to transfer overnight. After marking the wells, the membrane was air dried and cross-linked using a GS Gene Linker UV chamber (Bio-Rad Laboratories Inc., Mississauga, ON).

II.2.3.3.1 Hybridization of Southern blots

Hybridizations were done using a HI-12000 hybridization incubator (Tyler Research Instruments Corp., Edmonton, AB). The membrane was inserted into the hybridization bottle along with pre-hybridization solution [6X SSPE, 5X Denhardt's solution [2% (w/v) Ficoll 400, 2% (w/v) polyvinylpyrolidone and 20 mg/ml bovine serum albumin (57)], 0.1% SDS and 10 µg/ml denatured salmon sperm DNA], 0.125 ml per cm^2 of membrane. The 20X SSPE stock solution consists of 3 M NaCl, 0.2 M NaH_2PO_4 and 0.02 M EDTA pH 7.4. The membrane was incubated in the pre-hybridization solution for two hours. The hybridization began by replacing the pre-hybridization solution with an equal volume of fresh pre-hybridization solution and the denatured random primer labeled probe (see section II.4.5). Hybridizations were generally carried out at 65°C with constant rotation. The membrane was then washed at 65°C with 2X SSPE plus 0.1% SDS and 0.2X SSPE plus 0.1% SDS for 20 minutes each. After washing, the membrane was wrapped in saran wrap and exposed to Kodak X-OMAT AR film at -70°C for up to 7 days. Exposed X-ray films were developed using a Fuji RGII X-ray film processor (Fuji Photo Film Co., Tokyo, Japan).

When DNA from heterologous *Streptomyces* species was hybridized with fragments of DNA from the *S. clavuligerus* cephamycin cluster, the hybridizations were done at 55°C overnight. The next day the membrane was washed twice at 55°C with 2X SSPE plus 0.1% SDS for 20 minutes each.

If the membrane was going to be rehybridized, the membrane was added to 1 litre of boiling 0.1% SDS for 20 minutes.

II.2.4 *S. clavuligerus* fermentation culture conditions and media

Production cultures for measurement of cephamycin C production and/or production of cell-extracts for western analysis were prepared by inoculating spores of *S. clavuligerus* into 10 ml of trypticase soy broth supplemented with 1% (w/v) soluble starch (TSBS) to give a starting OD_{600} of 2.2×10^{-4} (based on the OD_{600} of a concentrated spore stock). These seed cultures were grown in universal bottles containing a stainless steel spring at 28°C on a rotary shaker (280 rpm) for 24 to 36 hours. If the seed cultures were started from plasmid containing strains or the NTG mutants, 20 µl of a mycelial glycerol stock was added as inoculum. The production cultures were prepared by subculturing sufficient seed culture into 25 ml of fresh TSBS in 125 ml flasks to give an OD_{600} of 8.0×10^{-2}, and incubating for 48 or 72 hours. The media in the production flasks were only supplemented with antibiotic if the strains carried autonomous plasmids. For characterization of the *lat* mutants, half of the production flasks were supplemented with 50 mM α-aminoadipic acid in 50 mM MOPS pH 7.5 to a final concentration of 2 mM.

Clavulanic acid or clavam production cultures were also prepared as described above except seed cultures were washed twice with 10 ml of sterile water before subculturing into 25 ml of fresh TSBS, starch asparagine (SA) medium or SOY medium and incubating for 72 hours. Starch asparagine medium consisted of 10 g soluble starch, 2 g asparagine, 21 g MOPS, 0.6 g $MgSO_4 \cdot 7H_2O$, 4.4 g K_2HPO_4 and 1 ml trace elements solution (0.1 g $FeSO_4 \cdot 7H_2O$, 0.1 g $MnCl_2 \cdot 4H_2O$, 0.1 g $ZnSO_4 \cdot 7H_20$ and 0.13 g $CaCl_2 \cdot 3H_2O$ per 100 ml water), pH 6.8 (4). SOY medium consisted of 15 g soybean flour, 4.7 g soluble starch, 0.1 g KH_2PO_4 and 0.2 g $FeSO_4 \cdot 7H_2O$, adjusted to pH 6.8 with NaOH (192).

For time course analysis of wildtype *S. clavuligerus*, spores were inoculated directly into the production flasks rather than using a seed culture. Spores of *S. clavuligerus* were added directly into 25 ml of TSBS in 125 ml flasks to give a starting OD_{600} of 5.5×10^{-4} (based on the OD_{600} of the concentrated spore stock). These flasks were harvested for analysis at the times designated by the experiment and during earlier time points multiple flasks were harvested.

II.2.4.1 Harvesting and preparation of cell-extracts

Using a clipped 1000 μl pipette tip, 0.5 ml and 1.0 ml aliquots of culture were transferred to 1.5 ml microfuge tubes and stored at -20°C. These samples were later used for bioassays, high performance liquid chromatography (HPLC) analysis or growth determination. If western analysis was also to be done, 10 ml of culture was pipetted into a centrifuge tube using a 10 ml wide bore pipette and centrifuged for 10 minutes. After the supernatant was decanted the mycelial pellets were prepared for disruption or they could be stored at -20°C but never for longer than a week. The pellets were washed with 5 ml of 0.85% (w/v) NaCl and centrifuged before being resuspended in 2 ml of TDE buffer (50 mM Tris pH 7.2, 0.01 mM EDTA and 0.1 mM DTT). The mycelia were then disrupted by sonication (2 x 20 seconds; micro-probe - power setting 2; Branson Sonifier 450; Branson Ultrasonic Corp., Danbury CT). The disrupted mycelia were transferred to a 1.5 ml centrifuge tube and microcentrifuged for 5 minutes at 4°C before the cleared cell-extract was decanted into fresh 1.5 ml tubes.

When the *lat* mutants were characterized the cultures being analyzed were harvested, washed and disrupted immediately and an aliquot of the cleared cell-extract was mixed with an equal volume of 2X SDS-PAGE sample buffer and stored at -20°C.

II.2.4.2 Analysis of cell-extracts for biosynthetic enzyme activities

II.2.4.2.1 δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase assay

The level of δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase (ACVS) activity in cell-extracts was measured by quantifying the condensation of the three precursor amino acids (cysteine, valine and α-aminoadipate) into the ACV tripeptide as described in Jensen et al. (116). Cell-extracts for ACVS assays were prepared in the ACVS-breakage buffer (5 mM $MgCl_2$, 1 mM EDTA, 30 mM β-mercaptoethanol, 50 mM KCl, 100 mM MOPS-NaOH pH 7.5 and 50% (v/v) glycerol).

A bulk reaction mix consisting of: 25 μl 40 mM disodium ATP, 3 μl of 1 M $MgCl_2$, 20 μl of 100 mM DTT, 50 μl of 100 mM phosphoenolpyruvate and 352 μl of 100 mM Tris-HCl pH 7.5 was made as was an amino acid reaction mix consisting of: 100 μl each of 50 mM L-valine, 50 mM L-α-aminoadipic acid and 10 mM L-cysteine. Each assay consisted of 45 μl of bulk reaction mixture, 30 μl of amino acid mixture and 25 μl of cell-extract and was incubated for 1 hour at room temperature. The assays were stopped by the addition of 25 μl of 20% (w/v) trichloroacetic acid, incubated on ice for 10 minutes and then microcentrifuged for 5 minutes.

A thiolyte derivatization mixture consisting of: 285 µl of 0.2 M N-[2-hydroxyethyl]piperazine-N'-3-propanesulfonic acid (EPPS)-KOH pH 8.0, 6 µl of 0.25 M diethylene-triamine-penta acetic acid and 9 µl of 0.1 M Thiolyte MB (Calbiochem, San Diego, CA) was prepared. Thirty microlitres of the thiolyte mixture was added with 10 µl of centrifuged ACVS assay mixture and incubated for 10 minutes in the dark before adding 5 µl of 1 M methanesulfonic acid to stop the derivatization and 10 µl of the derivatized ACVS assay solution was injected onto the HPLC. One unit of ACVS activity was defined as the amount of enzyme that produced 1 µmol of ACV per minute.

The chromatographic equipment used in the analysis included the Model 6000A and Model 45 HPLC pumps, a 712 Wisp sample processor, and a 420 fluorescence detector. Ten microlitre samples were injected onto a µbondpack C_{18} column and analyzed using a 30 minute gradient elution using buffer A: 0.25% NaOAC (adjusted to pH 5.0 with NaOH) plus 10% (v/v) methanol and buffer B: 0.25% NaOAC (adjusted to pH 5.0 with NaOH) plus 90% (v/v) methanol at a flow rate of 2 ml/minute. All HPLC equipment and columns were from Millipore Corp., Waters Chromatography (Milford, MA).

II.2.4.2.2 Isopenicillin N synthase bioassay

The level of isopenicillin N synthase (IPNS) activity in cell-extracts was estimated by quantitatively measuring the conversion of ACV to isopenicillin N as described in Jensen et al. (113). Isopenicillin N is the first bioactive molecule of the cephamycin pathway and the amount of isopenicillin N produced in the cell-free assay was quantitated using a standard curve for isopenicillin N versus the indicator organism *Micrococcus luteus*. One unit of IPNS activity was defined as the amount of enzyme that produced 1 µmol of isopenicillin N per minute.

A tube containing 100 µg of neutralized and lyophilized bis-ACV was resuspended in 25 µl of a 10 mg/ml stock of fresh DTT and incubated at room temperature for 15 minutes. The tube containing reduced ACV was supplemented with 25 µl of 11.2 mg/ml sodium ascorbate, 10 µl of 1.0 mg/ml $FeSO_4$ and 40 µl of 50 mM Tris-HCl pH 7.5 to create the reaction mix. A 10 µl aliquot of the reaction mix was added to 30 µl of cell-extract and incubated for 1 hour at room temperature. The reaction was stopped by the addition of 40 µl of methanol and chilled on ice. After microcentrifuging for 5 minutes, 20 µl samples were analyzed on *M. luteus* bioassay plates. A square bioassay plate was flooded with 100 ml of NBG agar (0.8% (w/v) nutrient broth, 0.1% (w/v) glucose and 2% (w/v) agar) into which 2 ml of an *M. luteus* glycerol stock was added. The plates were allowed to dry for approximately 30 minutes before 1/4 inch paper disks

(Schleicher & Schuell, Keene NH) were placed on the surface of the agar. Plates were incubated overnight at 28ºC and then zones of inhibition were measured.

II.2.4.2.3 Isopenicillin N epimerase bioassay

The level of isopenicillin N epimerase (IPNE) activity in cell-extracts was estimated by quantitatively measuring the conversion of isopenicillin N to penicillin N as described in Jensen et al. (114). The amount of penicillin N was quantitated using a standard curve for penicillin N versus the indicator organism *E. coli* ESS. One unit of IPNE activity was defined as the amount of enzyme that produced 1 µmol of penicillin N per minute.

A tube containing 20 µg of lyophilized isopenicillin N was resuspended with 50 µl of 100 mM Tris-HCl pH 7.0 and a 10 µl aliquot was added to 20 µl of cell-extract and incubated for 1 hour at room temperature. The reaction was stopped by the addition of 30 µl of methanol and chilled on ice. After microcentrifuging for 5 minutes, 20 µl samples were analyzed on the *E. coli* ESS bioassay plate (see II.2.4.4.1).

II.2.4.2.4 Desacetoxycephalosporin C synthase bioassay

The level of desacetoxycephalosporin C synthase (DAOCS) activity in cell-extracts was estimated by quantitatively measuring the conversion of penicillin N to desacetoxycephalosporin C as described in Jensen et al. (115). The amount of desacetoxycephalosporin C produced in the cell-free assay was quantitated using a standard curve for cephalosporin C versus the indicator organism *E. coli* ESS and supplementing the bioassay media with bacto-penase. One unit of DAOCS activity was defined as the amount of enzyme that produced 1 µmol of desacetoxycephalosporin C per minute.

A tube containing 20 µg of lyophilized penicillin N was resuspended with 12.5 µl of 10 mg/ml stock of fresh DTT, 12.5 µl of 11.2 mg/ml sodium ascorbate, 5 µl of 1.0 mg/ml $FeSO_4$, 15 µl of α-ketoglutarate mix (4.5 mg of α-ketoglutarate resuspended in 1.6 ml of 50 mM Tris-HCl pH 7.5, 0.2 ml of 1 M $MgSO_4$, 0.2 ml of 1 M KCl and 26 µl of 1 N NaOH) and 15 µl of 50 mM Tris-HCl pH 7.5. A 10 µl aliquot of this reaction mix was added to 5 µl of water and 5 µl of cell-extract and incubated for 1 hour at room temperature. The reaction was stopped by the addition of 20 µl of methanol and chilled on ice. After microcentrifuging for 5 minutes, 20 µl samples were analyzed on the *E. coli* ESS bioassay plate supplemented with 200 µl of bacto-penase (see II.2.4.4.1).

II.2.4.3 Western blot analysis of cell-extracts

II.2.4.3.1 Quantitation of protein content in cell-extracts

The protein concentration of each cell-extract was determined using the Bio-Rad protein assay (Bio-Rad Laboratories Inc.) based on the method of Bradford (30). A standard curve was constructed using bovine gamma globulin as a reference and stored into memory of the Unicam UV-3 spectrophotometer (Unicam Ltd., Cambridge, United Kingdom). A new standard curve was constructed for each new bottle of Bio-Rad protein assay stock solution.

II.2.4.3.2 SDS-PAGE for ACVS protein detection

The ACVS protein can be detected by SDS-PAGE because of its very large molecular weight it becomes separated from the majority of the other proteins in the cell-extract. The cell-extract was separated on a 10% SDS-PAG because the protein migrates as a nice sharp band very near the top of the gel. Analysis of cell-extracts on lower percentage gels allows the protein to migrate further into the gel but this causes the protein to become more diffuse and harder to detect.

Discontinuous 10% protein gels were cast and run using the Mini-Protean II gel apparatus (Bio-Rad Laboratories Inc.) as described in Blackshear (29). The separating gel was composed of: 10% (v/v) acrylamide (37.5:1.0), 0.2% (w/v) NaCl, 375 mM Tris-HCl pH 8.8 and 0.1% (w/v) SDS. After polymerization of the separating gel, the stacking gel [4.0% (v/v) acrylamide (37.5:1.0), 125 mM Tris-HCl pH 6.8 and 0.07% (w/v) SDS] was poured around 10 or 15 well combs. Both the separating and stacking gels were polymerized by the addition of ammonium persulfate and TEMED. The gels were electrophoresed using fresh SDS-PAGE running buffer (50 mM Tris-HCl pH 8.3, 384 mM glycine and 0.1% (w/v) SDS).

It has been determined that the ACVS protein is unstable in cell-extracts (7), like the LAT protein and therefore samples were immediately diluted in sample buffer for storage. Protein samples were prepared by adding cell-extract protein and water to an equal volume of 2X sample buffer (125 mM Tris-HCl 6.8, 4% (w/v) SDS, 10% (v/v) 2-mercaptoethanol, 0.5% (w/v) bromophenol blue and 20% (v/v) glycerol) and boiling for 5 minutes. Seventy five microgram amounts of cell-extract proteins were separated on the 10% SDS-PAGs. Gels were electrophoresed at 50 volts in the cold room until the bromophenol blue dye migrated to the bottom of the gel.

The gels were stained overnight in 0.5% (w/v) Coomassie blue plus 30% (v/v) methanol and 10% (v/v) acetic acid. The gels were destained in 50% (v/v) methanol and 10% (v/v) acetic acid and then 10% (v/v) acetic acid.

II.2.4.3.3 SDS-PAGE and transfer to PVDF membranes

Protein samples were prepared by adding cell-extract protein and water to an equal volume of 2X sample buffer and boiling for 5 minutes. Five or ten microgram amounts of cell-extract proteins were separated on the discontinuous 10% SDS-PAGs. Gels were electrophoresed at 50 to 150 volts until the bromophenol blue dye migrated to the bottom of the gel.

The two gels per apparatus were transferred to PVDF membranes using a Mini Transblot apparatus (Bio-Rad Laboratories Inc.) using the method of Towbin et al. (216). The gels were equilibrated for 15 minutes in transfer buffer (25 mM Tris-HCl pH 8.3, 150 mM glycine and 20% (v/v) methanol) and the PVDF membranes were first hydrated in methanol for 5 seconds before equilibrating in water for 5 minutes and then 15 minutes in transfer buffer. The transfer "sandwich" began with a moistened foam pad, a piece of wet filter paper, the protein gel, the PVDF membrane, another filter paper and the other foam pad. Transfer was done for 1 hour at 200 mAMPS per blotting apparatus in the cold room.

II.2.4.3.4 Developing western blots

Proteins were detected using the ECL chemiluminescent western system reagents and protocol manual (11) supplied with the kit. The primary antibodies were used at the following dilutions: purified LAT 1:400, CcaR 1:5000, LAT 1:7500, IPNS 1:10000 and DAOCS 1:10000 (see II.3.5.4 and II.3.5.5). DAOCS antibodies were generously provided by C. Reeves, Panlabs Inc. Donkey anti-rabbit immunoglobulin G horseradish peroxidase conjugate was used as the secondary antibody and at a 1:1000 dilution.

The transferred PVDF membranes were blocked in 5% (w/v) skim milk solution made in Tris buffered saline (20 mM Tris-HCl pH 7.6 and 137 mM NaCl) plus 0.1% (v/v) Tween 20 (TBST) either for 16 hours at 4°C or for one hour at room temperature. The blots were washed by gentle agitation with TBST plus 0.5% (w/v) skim milk (TBSTM); two quick rinses, one 15 minute wash followed by two 5 minute washes. The primary antibody was added to fresh TBSTM and incubated for 60 minutes. The blots were again washed with TBSTM; two quick rinses, one 15 minute wash followed by two 5 minute washes. The secondary antibody (1:1000) was added to fresh TBSTM and incubated for 60 minutes. The blots were again washed with TBSTM; two quick rinses, one 15 minute wash followed by three 5 minute washes.

The blots were developed by floating the membrane in 2 ml of detection reagent (equal volumes of detection reagents 1 and 2) for 1 minute before sealing in saran wrap.

The developed blots were exposed immediately for up to 10 minutes on Fuji RX X-ray film.

Western blots were stripped for reprobing by incubating the membranes in 100 mM 2-mercaptoethanol, 2% (w/v) SDS and 62.5 mM Tris pH 6.7 at 50°C for 30 minutes. Blots were rinsed well with TBST before reblocking in the 5% (w/v) skim milk solution.

II.2.4.4 Quantitation of antibiotic production and growth

II.2.4.4.1 Bioassay for penicillin/cephamycin

Total penicillin/cephamycin antibiotic production by *S. clavuligerus* was assayed using the agar-diffusion method with the indicator organism *E. coli* ESS (112). A square (22.5 cm by 22.5 cm) bioassay plate (Nunc Laboratories, Denmark) was flooded with 100 ml of TCS agar (3% (w/v) TSB and 2% (w/v) agar) into which 2 ml of an *E. coli* ESS glycerol stock was added. Plates were then allowed to dry for approximately 30 minutes before the 0.5 ml aliquots of *S. clavuligerus* culture were thawed rapidly in a water bath and centrifuged for 2 minutes to pellet the mycelia. The 1/4 inch paper disks were placed on the surface of the agar and spotted with 20 µl of culture supernatant in triplicate. Plates were incubated overnight at 37°C before the zones of inhibition were measured.

For the characterization of the *lat* mutants cephamycin production values were calculated as specific activities relative to the OD_{595} calculated for each culture. Whereas during the characterization of the *ccaR* mutants cephamycin production values were calculated volumetrically.

II.2.4.4.2 HPLC analysis for clavulanic acid and other clavams

The production of clavulanic acid and the clavams by *S. clavuligerus* fermentations in SA or SOY medium was assayed by HPLC (80, 168). The culture supernatants were microcentrifuged for 5 minutes to remove mycelia and 100 µl of cleared supernatant or a suitable dilution in water, was derivatized with 25 µl of 25% (w/v) imidazole pH 6.8 for 15 minutes. The samples were microcentrifuged again for 5 minutes and 100 µl was then transferred to HPLC sample tubes. Non-derivatized controls, in which 25 µl of water replaced the imidazole in the derivatization reaction, were used to remove background absorbance by the sample.

Fifty microlitre samples were injected onto a µbondpack C_{18} column and analyzed using a 15 minute isocratic elution using a 100 mM NaH_2PO_4 (adjusted to pH 3.68 with H_3PO_4) plus 10% methanol buffer at a flow rate of 2 ml/minute. The 480 programmable multi-wavelength detector was set to detect at 311 nm. The peaks were identified and quantitated next to chromatograms of authentic standards of clavulanic acid and clavam-2-carboxylate.

II.2.4.4.3 Growth determination

A method was developed by Malmberg et al. (143) in which the clumping nature of the *Streptomyces* can be overcome to allow accurate measurements of culture optical densities.

A 0.5 ml sample of TSBS-grown culture was added to 3 ml of water and 0.5 ml of 2.5 M HCl in a conical centrifuge tube. The sample was sonicated for 30 seconds at setting 1 to give a homogeneous mixture. The optical density of the suspension was measured in triplicate after diluting 5-fold in 0.25 M HCl to give an OD_{595} value less than 0.6. Samples in which no antibiotic was produced only a single reading was determined.

II.2.4.5 Harvesting cultures for *xylE* promoter probe analysis

A 10 ml sample of culture was pipetted into a centrifuge tube and centrifuged for 10 minutes. The supernatants were decanted and the pellets were washed with 5 ml of 20 mM KPO_4 buffer pH 7.2 and centrifuged. Again the supernatant was discarded and the pellets were resuspended in 1.5 ml of breakage buffer [100 mM KPO_4 pH 7.5, 20 mM EDTA and 10% (v/v) acetone (100)]. The mycelial suspension was then disrupted by sonication (2 x 20 seconds; micro-probe - power setting 2). Then 1.5 ml of the disrupted mycelia was transferred to 1.5 ml centrifuge tubes and microcentrifuged for 10 minutes at 4°C before the cleared cell-extract was decanted into fresh 1.5 ml tubes.

II.2.4.5.1 Catechol 2,3 dioxygenase assays

The cell-extracts for the catechol 2,3 dioxygenase (C23O) assays could not be frozen so they were stored on ice and assayed immediately. The protein content of the cell-extracts was determined later from the assayed samples stored at -20°C.

The assay used was a modification of the C23O assay designed by Zukowski et al. (237). The light-sensitive catechol substrate was made as a 20 mM stock in ethanol and stored in the dark at 4°C. The C23O activity was measured as a change of absorbance at 375 nm per minute and this was later converted to specific activity (units of activity x 10^{-6} per mg of protein). Under the conditions of this assay the molar absorption coefficient of the C23O product 2-hydroxymuconic semialdehyde was reported to be 3.3×10^4 (190). One unit of activity is defined as the formation of 1 μmol of 2-hydroxymuconic semialdehyde per minute.

The assay was started by the addition of up to 35 µl of enzyme to a test tube containing 5 µl of 20 mM catechol and 960 µl of 100 mM KPO_4 buffer pH 6.8. The remainder of the 35 µl volume was made up with breakage buffer. The reactions were allowed to incubate for 1 minute, during which the reaction mixes were being transferred to 1 ml cuvettes and placed in the spectrophotometer, before the change in A_{375} was followed for the next 4 minutes. Up to four samples were monitored at one time. Samples were incubated at ambient temperature rather than 30ºC.

II.2.5 Generation of gene replacement mutants by homologous recombination

Double crossover mutants were generated by transforming protoplasts of *S. clavuligerus* with targeting vectors carrying various mutated versions of the genes of interest. The targeting vectors are *E. coli/Streptomyces* shuttle plasmids using the pIJ101 replicon-containing plasmids pJOE829 and pIJ486 for replication in *Streptomyces*.

Cell material from the patches of two primary *S. clavuligerus* transformants were used to inoculate TSBM medium plus antibiotics selecting for both the vector and the disruption markers. These cultures were grown for approximately 48 hours before they were bioassayed for antibiotic production, used to prepare glycerol stocks, streaked onto ISP3 without antibiotics and used to isolate plasmid DNA. The primary transformants were checked to ensure that they did produce antibiotic (cephamycin) before beginning the gene replacement experiment and the isolated plasmid DNA was transformed into *E. coli* to check for rearrangement. The primary transformants were streaked onto ISP3 without antibiotic selection because of the segregational instability of the pIJ101 replicon in *S. clavuligerus* during sporulation. This leads to a loss of plasmid during sporulation but also at a variable frequency the plasmid crosses over into the chromosome via the mutated homologous DNA insert and then crosses back out to complete the double crossover.

The protocol for the gene disruption experiments was followed as outlined in Aidoo et al. (5). The filtered spores obtained from the non-selective sporulation plates were serially diluted in 0.1% (w/v) peptone water and 50 to 200 µl was plated out onto ISP3 plates and incubated at 28ºC until the individual colonies have sporulated again. Two plates of spores for each primary transformant, containing between 150 to 300 colonies, were replica plated sequentially onto: MYM agar plates plus antibiotic for the disrupting marker, MYM plates plus antibiotic for the vector marker and MYM plates. Putative disruption mutant colonies were recognized by their maintenance of resistance to the disrupting marker and sensitivity to the vector marker. These putative disruptants were patched onto MYM plates without antibiotics before repatching onto MYM plates containing the vector marker antibiotic or the disruption marker antibiotic to confirm the antibiotic resistance phenotypes of the putative mutants.

For one gene replacement experiment the goal was not to replace the wildtype *lat* gene with a mutant gene containing an antibiotic resistance marker but rather to remove the resistance marker from a mutant strain but to still leave behind a mutated *lat* gene. When the pDA565 and pDA566 repair plasmids were used to remove the *apr* marker from the *lat::apr* mutant the putative disruptants were screened for the loss of the plasmid resistance marker as well as the loss of *apr* marker from the strain.

Two mutants for each gene replacement experiment were chosen if both orientations of disrupting/deleting marker were present. If only one orientation was present then four different mutant clones were chosen. If possible, mutants were picked from both of the primary transformants.

II.3 *E. coli* procedures

All manipulations involving *E. coli* cultures were done aseptically. All centrifugation procedures were done using a JA-20 rotor in an Beckman centrifuge at 5000 rpm and cooled to 4ºC, unless otherwise noted. All microcentrifugation steps were done in an Eppendorf 5415 C microcentrifuge (Brinkmann Instruments Inc. Mississauga ON) at full speed (14000 rpm), unless otherwise noted.

II.3.1 Storage of strains

E. coli strains with or without plasmids were maintained as glycerol stocks generated by combining equal volumes of culture and 50% (w/v) glycerol and storing at -70ºC. A list of the bacterial strains used or created in this study is provided in Table 2.

II.3.2 Culture conditions and media

E. coli cultures were grown at 37ºC in 2x YT to make competent or electrocompetent cells or for protein expression experiments. Cultures were grown in Terrific Broth plus 0.5% (w/v) glucose and antibiotic for plasmid isolation. Recipes for terrific broth and 2x YT were as described (194). Transformations were plated out onto 2x YT agar plates plus 0.5% glucose and antibiotic. In transformation experiments using α-complementation for visual selection of recombinant plasmids both isopropyl β-D-thiogalactopyranoside (IPTG) and 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (X-gal) were supplemented into 2x YT agar plates without glucose as described in Sambrook et al. (194). When kan was used as the selectable marker plates were composed of trypticase soy broth agar plus 0.5% glucose and kan. Apr and amp were used at final concentrations of 100 µg/ml while kan and hyg were used at final concentrations of 50 µg/ml.

II.3.3 Introduction of DNA into *E. coli*

II.3.3.1 Preparation of electrocompetent *E. coli* cells

Electrocompetent *E. coli* were prepared as a modification of the method in Sambrook et al. (194). Five millilitres of 2x YT were inoculated with 5 µl of a glycerol stock of the desired *E. coli* strain and incubated overnight on a tube roller. The overnight culture was used as a 1% (v/v) inoculum for fresh 2x YT in a flask, with shaking at 250 rpm, and the culture was grown for approximately 2 hours until the OD_{600} was between 0.5 and 0.9 before centrifugation for 10 minutes. The supernatant was poured off and the cells were washed with an equal volume of ice-cold 0.1 M glucose. The washed cells were again centrifuged and the wash was repeated with a half volume of 0.1 M glucose. The wash was poured off and the pellet was resuspended to 1/100 of the original culture volume with 15% (w/v) glycerol. Then 40 µl of cells were aliquoted into 1.5 ml centrifuge tubes and rapidly frozen on dry ice before storing at -70ºC.

II.3.3.2 Electroporation of DNA into electrocompetent *E. coli*

Ligation reaction mixtures were ethanol precipitated and resuspended in milliQ water to reduce the conductivity of the DNA sample before transformation. An aliquot of the ligation mixture was added to the electrocompetent cells which had been thawed on ice and then the entire mixture was transferred to a 2 mm gap electroporation cuvette for use with a Gene Pulser II electroporation apparatus (Bio-Rad Laboratories Ltd.). The cells/DNA mixture were pulsed at a setting of 25 µF capacitance, 200 ohms resistance and 2.5 kV and then immediately after the pulse 800 µl of 2x YT was added and the cells were transferred into a 1.5 ml centrifuge tube and incubated at 37ºC to recover. The recovery time for most transformations was 1 hour but for transformations where apr was used as the selectable marker more consistent results were obtained if the recovery period was extended to 2 hours.

II.3.3.3 Preparation of $MnCl_2$ competent *E. coli* cells

$MnCl_2$ competent *E. coli* cells were prepared using the method of S. Tiong and D. Nash (213). Five millilitres of 2x YT were inoculated with 5 µl of a glycerol stock of the desired *E. coli* strain and incubated overnight on a tube roller. The overnight culture was subcultured at 1% (v/v) into 200 ml of fresh 2x YT in a flask and grown for approximately 2 hours with shaking at 250 rpm until the OD_{600} was between 0.45 and

0.7. The culture was then centrifuged at 3000 rpm for 10 minutes. The cell pellets were resuspended in 40 ml of ice-cold, filter-sterilized solution A (30 mM KOAc, 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$ and 15 % (w/v) glycerol) and incubated on ice for 1 hour. The cell suspensions were pooled and centrifuged for 10 minutes before the pellets were resuspended in 12 ml ice-cold, filter-sterilized solution B (10 mM MOPS pH 7.0, 75 mM $CaCl_2$, 10 mM KCl and 15 % (w/v) glycerol). Then 200 or 500 µl of cells were aliquoted into 1.5 ml centrifuge tubes and rapidly frozen on dry ice before storing at -70°C.

II.3.3.4 Transformation of DNA into $MnCl_2$ competent *E. coli*

The ligation reaction mixture was added to 200 µl of the competent cells and incubated on ice for 30 minutes. The cells/DNA mixture was heat shocked at 42°C for 90 seconds, returned to ice for 90 seconds before 800 µl of recovery medium (2x YT + 20 mM glucose + 100 mM $MgCl_2$) was added and the transformation(s) incubated at 37°C. The recovery time for most transformations was 1 hour but for transformations where apr was used as the selectable marker more consistent results were obtained if the recovery period was extended to 2 hours.

If plasmid DNA isolated from *E. coli* was being transformed into another strain of *E. coli*, 2 µl of plasmid DNA was mixed with 20 µl of cells and only 200 µl of recovery medium was added after heat shocking. If plasmid DNA isolated from *Streptomyces* was being transformed into *E. coli*, 5 µl of plasmid DNA was mixed with 50 µl of cells and only 200 µl of recovery medium was added after heat shocking.

II.3.4 Analysis of DNA from *E. coli*

II.3.4.1 Plasmid DNA isolation

Isolation of *E. coli* plasmids was accomplished using a modification of the Birnboim and Doly (28) method listed in Sambook et al. (194). When ultrapure and/or highly supercoiled plasmid DNA was needed for sequencing or exoIII deletion experiments, Wizard miniprep columns were used for plasmid isolation according to the manufacturers' instructions.

II.3.4.2 Southern blotting and hybridization

Procedures were the same as those described earlier for use with *Streptomyces* (See II.2.3.4 and II.2.3.5).

II.3.4.3 Colony hybridization

Colonies from the transformation plate were replica plated onto two fresh 2x YT plus amp plates one of which had a Hybond N circle on the surface of the medium and the other plate served as the master plate. The plates were incubated for 12 hours and then the filter was stripped off the plate and sequentially incubated, colony side up, for 5 minutes in dishes containing filter paper soaked in: 1) 10% (w/v) SDS, 2) 0.5 M NaOH, 1.5 M NaCl, 3) 0.5 M Tris pH 7.4, 1.5 M NaCl and 4) 2 X SSPE (194). After the filter was air dried and UV cross-linked it was placed in 2X SSPE and scrubbed with a kimwipe to remove the leftover colony debris. The colony blots were hybridized in the same manner as Southern blots (see II.2.3.5).

II.3.4.4 DNA sequence analysis

DNA sequence analyses were conducted using template DNA prepared with the Wizard miniprep columns. Template DNA was ethanol precipitated and resuspended in 20 μl of water. This DNA was quantified by using a TKO 100 DNA fluorometer (Hoefer Scientific Instruments, San Francisco, CA) as 1 μg of DNA was needed for each reaction. Double stranded DNA cycle sequencing was carried out by the Department of Biological Sciences DNA Synthesis Laboratory (University of Alberta). The reverse, forward, T3 and T7 primers were typically used to prime the sequencing reactions but custom oligonucleotide primers were also employed.

DNA Strider 1.2 (144) was used to store sequencing data and to search for restriction sites present within the sequenced DNA. Open reading frames (ORFs) were detected by %G+C bias present in *Streptomyces* (25) using the codonpreference program within the UWGCG package (Genetics Computer Group, Inc., Madison, WI; 60). BLAST (Basic Local Alignment Search Tool) homology searches (9, 10) were done over the internet at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/). The percentage of identical amino acids present in two homologous ORFs was measured using the LFASTA alignment program (172) available over the internet (http://dot.imgen.bcm.tmc.edu:9331/seq-search/alignment.html). The multiple alignment of homologous ORFs was achieved using the Clustal W1.7 multiple alignment program (212) available over the internet (http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html). The presence of conserved or similar amino acids in the Clustal W1.7 alignment was illustrated using the Boxshade 3.21 program available over the internet (http:ulrec3.unil.ch/software/BOX_form.html). The position of the signal sequence processing sites was predicted using the SignalP program (160) available over the internet (http://www.cbs.dtu.dk/services/SignalP/# submission).

II.3.5 Production of MBP-fusion proteins with LAT or CcaR in *E. coli* to create polyclonal antibodies in rabbits

The *ccaR* and *lat* genes were fused with the *malE* gene in the vector pMAL-c2. These recombinant vectors would allow the production of the CcaR or LAT protein fused to the C-terminal end of the maltose binding protein (MBP). Artificial *Bam*HI sites were generated on the 5' end of each gene just before the start codon. This created a *ccaR* or *lat* gene capable of creating a translational fusion with the MBP protein when inserted into pMAL-c2.

Amplification, cloning and sequencing of the *ccaR* and *lat* genes was described in II.4.8.1 and II.4.8.2, respectively. The resulting expression plasmids, pMAL-CcaR and pMAL-LAT, are described in Table 3.

II.3.5.1 Production of MBP-fusion proteins

Overnight cultures of pMAL-CcaR and pMAL-LAT were grown in 2x YT before diluting 1:100 into 50 ml of prewarmed 2x YT and incubating for 1 hour at 37°C. Cultures were induced by adding IPTG to a final concentration of 0.5 mM and then incubated at 30°C for 4 hours.

II.3.5.2 Purification of MBP-fusion proteins

The MBP-fusion proteins were purified from induced cultures as described in the New England Biolabs manual (156) that came with the plasmids. Cultures were harvested by centrifugation for 10 minutes, washed and then the cell pellets were resuspended in 1 ml of STE buffer (200 mM NaCl, 1 mM EDTA and 20 mM Tris pH 7.4). Cells were broken by sonication (3 x 20 seconds; micro-probe - power setting 1) and then centrifuged to remove unbroken cells and inclusion bodies. The supernatant was added to a tube containing 1 ml of amylose resin equilibrated in STE buffer and incubated on ice for 15 minutes before washing 6 times with 5 ml of STE. The MBP fusions were eluted by twice adding 1 ml of STE buffer plus 10 mM maltose.

The MBP-fusion proteins were not treated with Factor Xa protease after purification because of the expense of the protease and the number of steps that would be required to purify away the MBP and the protease.

Table 3. Plasmids list.

Plasmid	Description
Cloning vectors	
pBluescript KS/SK	General cloning and generation of exoIII deletions. Amp^R. KS(+) or SK(+) with MCS flanked by the reverse or T3 and the -21 or T7 primer binding sites. (Stratagene cloning systems, La Jolla, CA)
pKSN/SKN	Derivatives of pKS(+) or pSK(+) in which the *Bam*HI site has been mutated by insertion of the Apa8 adapter oligonucleotide which creates an *Nco*I site. Amp^R. (This study)
pTZ18R	General cloning and generation of exoIII deletions. Amp^R. MCS flanked by the reverse and -21 primer binding sites. Very good for cloning large fragments of DNA. (United States Biochemical Corp., Cleveland OH)
pUC118/119	General cloning. Amp^R. MCS flanked by the reverse and -21 primer binding sites. From P. Vieira. (219)
pUC8/9	General cloning. Amp^R. MCS flanked by the reverse and -21 primer binding sites. From P. Vieira. (149)
pIJ2925	General cloning. Amp^R. MCS of pUC18 flanked by *Bgl*II sites. From M. Bibb. (104)
pSL1180	General cloning. Amp^R. Contains "superlinker" MCS which has many unique MCS enzyme sites. (Pharmacia Biotech, Baie d' Urfé, QC)
pMAL-c2	*E. coli* IPTG inducible expression plasmid. High copy number expression vector creating a translational fusion between the MBP and the cloned gene. Amp^R. (New England Biolabs Ltd., Mississauga, ON)
pGEX-2T	*E. coli* IPTG inducible expression plasmid. High copy number expression vector creating a translational fusion between the GST protein and the cloned gene. Amp^R. (Pharmacia Biotech)
pFD666	*E. coli/Streptomyces* shuttle plasmid. High copy number cloning vector with small MCS in *E. coli*. Bifunctional Kan^R/Neo^R. Mid-copy number in *Streptomyces*. From R. Brzezinski. (59)
pUWL-KS/SK	*E. coli/Streptomyces* shuttle plasmid. High copy number cloning vector with pBluescript KS(+) or SK(+) MCS in *E. coli*. Amp^R. Tsr^R. High copy number plasmid for *Streptomyces* with the pIJ101 replicon. From U. Wehmeier. (224)
pSET152	*E. coli/Streptomyces* shuttle plasmid. High copy number cloning vector with small MCS in *E. coli*. Bifunctional Apr^R. Non-replicative in *Streptomyces* but capable of site-specific chromosomal integration at the ØC31 attachment site. From NRRL. (27)

Table 3. (continued)

Plasmid	Description
Cloning vectors	
pHM8a	*E. coli/Streptomyces* expression shuttle plasmid. High copy number cloning vector with small MCS following the *ermE** promoter and designed *Streptomyces* S/D in *E. coli*. Bifunctional Hyg^R. Non-replicative in *Streptomyces* but capable of site-specific chromosomal integration at the mini-circle attachment site. From H. Motamedi. (153)
pJOE829	High copy number plasmid for *Streptomyces* with the pIJ101 replicon. Bifunctional Hyg^R. Small MCS for general cloning. From J. Altenbuchner.
pIJ4083	High copy number promoter probe plasmid for *Streptomyces* with the pIJ101 replicon. Tsr^R. A small MCS precedes the *xylE* reporter gene both of which are flanked by transcription terminators. From M. Bibb. (47)
pIJ486	High copy number promoter probe plasmid for *Streptomyces* with the pIJ101 replicon. Tsr^R. A small MCS precedes the *aph* reporter gene for cloning. From D. Hopwood. (223)
Subcloning cephamycin cluster fragments	
pOW309	*E. coli/Streptomyces* shuttle cosmid containing most of the cephamycin cluster spanning from the end of the *rhsA* gene to midway through the *pcbAB* gene. Bifunctional Apr^R and Neo^R. From P. Skatrud. (185)
pOW309 Δ*Bam*HI	Derivative of the pOW309 shuttle cosmid created by self-ligation after *Bam*HI digestion. Removed most of the cephamycin cluster insert, leaving only parts of the *rhsA* and *pcbAB* genes. (This study)
pOW309 Δ*Kpn*I	Derivative of the pOW309 shuttle cosmid created by self-ligation after *Kpn*I digestion. Removed some of the cephamycin cluster insert, leaving from the *rhsA* gene to the *pcd* gene as well as part of the *pcbAB* gene. (This study)
pLAFR/2-49	*E. coli* cosmid containing some of the cephamycin cluster. Contains DNA from at least part of the *rhsA* gene to the *cefE* gene. Tc^R. (Jensen, this laboratory)
pDA100	8.3 kb *Bam*HI fragment B1 from pOW309 inserted into pKS(+) *Bam*HI. Contains the *ccaR*, *orf11*, *blp* and *lat* genes and part of the *pcbAB* gene. (This study)
pDA102	6.4 kb *Bam*HI fragment B2 from pOW309 inserted into pKS(+) *Bam*HI. Contains the *cmcH*, *cmcI*, *cmcJ* and *cefF* genes and part of the *cefD* gene. (This study)

Table 3. (continued)

Plasmid	Description
Subcloning cephamycin cluster fragments	
pDA103	6.4 kb *Bam*HI fragment B2 from pOW309 inserted into pKS(+) *Bam*HI in the opposite orientation relative to pDA102. Contains the *cmcH*, *cmcI*, *cmcJ* and *cefF* genes and part of the *cefD* gene. (This study)
pDA106	4.8 kb *Bam*HI fragment B3 from pOW309 inserted into pTZ18R *Bam*HI. Contains part of the *pcbAB* gene. (This study)
pDA107	4.1 kb *Bam*HI fragment B4 from pOW309 inserted into pKS(+) *Bam*HI. Contains the *cmcT* gene and part of the *pbpA* gene. (This study)
pDA109	3.2 kb *Bam*HI fragment B5 from pOW309 inserted into pSK(+) *Bam*HI. Contains the *cefE* and *pcd* genes and part of the *cefD* gene. (This study)
pDA110	3.2 kb *Bam*HI fragment B5 from pOW309 inserted into pSK(+) *Bam*HI in the opposite orientation relative to pDA109. Contains the *cefE* and *pcd* genes and part of the *cefD* gene. (This study)
pDA111	1.7 kb *Bam*HI fragment B6 from pOW309 inserted into pKS(+) *Bam*HI. Contains the *bla* gene and part of the *pbpA* and *sclU* genes. (This study)
pDA113	3.9 kb *Eco*RI/*Bam*HI fragment EB1 from pOW309 Δ*Bam*HI inserted into pKS(+) *Eco*RI/*Bam*HI. Contains part of the *pcbAB* gene. (This study)
pDA115	0.8 kb *Eco*RI/*Bam*HI fragment EB2 from pOW309 Δ*Bam*HI inserted into pKS(+) *Eco*RI/*Bam*HI. Contains part of the *rhsA* and *sclU* genes. (This study)
pDA116	0.8 kb *Eco*RI/*Bam*HI fragment EB2 from pOW309 Δ*Bam*HI inserted into pSK(+) *Eco*RI/*Bam*HI. Contains part of the *rhsA* and *sclU* genes. (This study)
pDA119	2.8 kb *Kpn*I/*Bam*HI fragment KB1 from pLAFR/2-49 inserted into pTZ18R *Kpn*I/*Bam*HI. Contains part of the *rhsA* and *sclU* genes. (This study)
pDA120	3.7 kb *Eco*RI/*Bam*HI fragment from pDA100 inserted into pKS(+) *Eco*RI/*Bam*HI. Contains the *ccaR*, *orf11* and *blp* genes. (This study)
pDA121	3.7 kb *Eco*RI/*Bam*HI fragment from pDA100 inserted into pSK(+) *Eco*RI/*Bam*HI. Contains the *ccaR*, *orf11* and *blp* genes. (This study)

Table 3. (continued)

Plasmid	Description
Subcloning cephamycin cluster fragments	
pDA122	1.7 kb *Bam*HI/*Bcl*I internal deletion of pDA120. Leaves behind a 2.0 kb *Bcl*I/*Eco*RI fragment that contains the *blp* gene and part of the *orf11* gene. (This study)
pDA123	1.7 kb *Bam*HI/*Bcl*I internal deletion of pDA121. Leaves behind a 2.0 kb *Bcl*I/*Eco*RI fragment that contains the *blp* gene and part of the *orf11* gene. Stable only on media supplemented with 0.5% (w/v) glucose. (This study)
pDA124	0.3 kb deletion of pDA107 created by self-ligation of a blunt-ended internal *Sph*I site and the MCS *Eco*ICRI. Contains the *cmcT* gene and part of the *pbpA* gene and was created for sequencing. (This study)
pDA125	1.4 kb deletion of pDA111 created by self-ligation of a blunt-ended internal *Bcl*I site and the MCS *Eco*RV. Contains the *bla* gene upstream sequences and was created for sequencing. (This study)
pDA126	1.4 kb deletion of pDA111 created by self-ligation of a blunt-ended internal *Nar*I site and the MCS *Eco*ICRI. Contains the *bla* gene downstream sequences and was created for sequencing. (This study)
pDA128	3.2 kb *Nco*I fragment from pDA100 inserted into pSKN *Nco*I. Contains the *blp* gene and part of the *orf11* and *lat* genes. (This study)
pDA130	4.7 kb *Nco*I fragment from pOW309 inserted into pSKN *Nco*I. Contains the *cmcH* and *ccaR* genes as well as part of the *cefF* and *orf11* genes. (This study)
pDA131	4.3 kb *Eco*RI/*Bam*HI fragment from pDA100 inserted into pKS(+) *Eco*RI/*Bam*HI. Contains the *lat* gene and part of the *pcbAB* gene. (This study)
pDA133	3.7 kb *Eco*RI/*Bam*HI fragment from pDA120 inserted into pTZ18R *Eco*RI/*Bam*HI. Contains the *ccaR*, *orf11* and *blp* genes. (This study)
pDA134	Destruction of the unique *Pst*I site in the MCS of pDA133 by digestion with *Pst*I, fragment ends were made blunt with T4 DNA polymerase and self-ligated. (This study)
pDA140	1.5 kb internal deletion removing the *apr* marker from pDA522 by *Pst*I digestion and self-ligated. Contains a mutant *ccaR* gene with the TSF at the *Eco*ICRI site and wildtype *orf11* and *blp* genes. (This study)

Table 3. (continued)

Plasmid	Description
Subcloning cephamycin cluster fragments	
pDA141	1.5 kb internal deletion removing the *apr* marker from pDA524 by *Pst*I digestion and self-ligation. Contains a mutant *orf11* gene with the TSF at the blunt-ended *Nco*I site and wildtype *ccaR* and *blp* genes. (This study)
pDA142	1.5 kb internal deletion removing the *apr* marker from pDA526 by *Pst*I digestion and self-ligation. Contains a mutant *blp* gene with the TSF at the blunt-ended *Not*I site and wildtype *ccaR* and *orf11* genes. (This study)
pDA143	1.5 kb internal deletion removing the *apr* marker from pDA520 by *Pst*I digestion and self-ligation. Contains a mutant *ccaR* gene and wildtype *orf11* and *blp* genes. (This study)
pDA150	1.7 kb *Bam*HI/*Bcl*I fragment from pDA120 inserted into pSL1180 *Bam*HI/*Bcl*I. Contains the *ccaR* gene and part of the *orf11* gene. (This study)
pIJ4070	380 bp *ermE** promoter fragment in pIJ2925. From M. Bibb. (26)
pDA151	0.4 kb *Eco*RI/*Bam*HI fragment from pIJ4070 inserted into pDA150 *Eco*RI/*Bam*HI. Contains the *ccaR* gene and also part of the *orf11* gene with the *ermE** promoter placed in tandem with the *ccaR* promoter. (This study)
pDA152	1.4 kb *Bam*HI/*Nru*I fragment from pDA120 inserted into pSK(+) *Bam*HI/*Eco*RV. Contains the *ccaR* gene. (This study)
pDA153	1.4 kb *Bam*HI/*Nru*I fragment from *S. clavuligerus dclX*68 genomic DNA inserted into pSK(+) *Bam*HI/*Eco*RV. Contains the *dclX*68 *ccaR* gene. (This study)
pDA154	1.4 kb *Bam*HI/*Nru*I fragment from *S. clavuligerus* NTG 26 genomic DNA inserted into pSK(+) *Bam*HI/*Eco*RV. Contains the NTG 26 *ccaR* gene. (This study)
pDA155	1.4 kb *Bam*HI/*Nru*I fragment from *S. clavuligerus* NTG 32 genomic DNA inserted into pSK(+) *Bam*HI/*Eco*RV. Contains the NTG 32 *ccaR* gene. (This study)
pDA156	1.4 kb *Bam*HI/*Nru*I fragment from *S. clavuligerus* NTG 83 genomic DNA inserted into pSK(+) *Bam*HI/*Eco*RV. Contains the NTG 83 *ccaR* gene. (This study)
pDA160	0.9 kb deletion generated by the self-ligation of pDA150 digested with *Sac*I. Contains the *ccaR* promoter and some of the *ccaR* gene. (This study)

Table 3. (continued)

Plasmid	Description
Subcloning cephamycin cluster fragments	
pDA161	0.8 kb deletion generated by the self-ligation of pDA150 digested with *Stu*I. Contains the *ccaR* promoter and most of the *ccaR* gene. (This study)
pDA162	0.7 kb deletion generated by the self-ligation of pDA150 digested with *Kpn*I. Contains the *ccaR* promoter and most of the *ccaR* gene. (This study)
pDA163	0.7 kb deletion generated by exoIII deletion of pDA150. Contains the *ccaR* promoter and most of the *ccaR* gene, the *ccaR* gene ends at position 19166 in Fig. 8. (This study)
pDA164	0.7 kb deletion generated by exoIII deletion of pDA150. Contains the *ccaR* promoter and most of the *ccaR* gene, the *ccaR* gene ends at position 19189 in Fig. 8. (This study)
pDA170	3.2 kb *Apa*I fragment from pDA100 inserted into pKS(+) *Apa*I. Contains the *blp* and *lat* genes and part of the *orf11* gene. (This study)
pDA171	3.2 kb *Apa*I fragment from pDA100 inserted into pKS(+) *Apa*I. Contains the *blp* and *lat* genes and part of the *orf11* gene in the opposite orientation. (This study)
pDA172	1.8 kb *Eco*RI/*Kpn*I (the *Kpn*I digestion was a partial) fragment from pDA131 inserted into pUC119 *Eco*RI/*Kpn*I. Contains the entire *lat* gene. (This study)
pDA173	1.6 kb *Eco*RI/*Eco*RV fragment from pDA171 inserted into pTZ18R *Eco*RI/*Eco*ICRI. Contains the entire *lat* gene. (This study)
pDA174	0.7 kb deletion generated by the self-ligation of pDA172 digested with *Kpn*I. Contains the *lat* promoter and part of the *lat* gene. (This study)
pDA175	225 bp *Eco*RI/*Nco*I fragment from pDA172 inserted into pKSN *Eco*RI/*Nco*I. Contains the *lat* promoter. (This study)
pDA177	1.6 kb *Nco*I/*Xba*I fragment (the *Nco*I end was made blunt with Klenow polymerase) from pDA172 inserted into pDA513 *Nde*I/*Xba*I (the *Nde*I end was made blunt with Klenow polymerase). Contains the *lat* coding region fused to the *ermE** promoter and S/D. (This study)
MBP or GST fusion constructs	
pDA180	0.8 kb *Eco*RI/*Bam*HI PCR fragment inserted into pKS(+) *Eco*RI/*Bam*HI. Contains the *ccaR* coding region. (This study)

Table 3. (continued)

Plasmid	Description
MBP or GST fusion constructs	
pMAL-CcaR	0.8 kb *Bam*HI/*Hin*dIII fragment from pDA180 inserted into pMAL-c2 *Bam*HI/*Hin*dIII. Contains the *ccaR* coding region in a translational fusion with MBP. (This study)
pGEX-CcaR	0.8 kb *Bam*HI/*Eco*RI fragment from pDA180 inserted into PGEX-2T *Bam*HI/*Eco*RI. Contains the *ccaR* coding region in a translational fusion with GST. (This study)
pDA181	1.4 kb *Eco*RI/*Bam*HI PCR fragment inserted into pKS(+) *Eco*RI/*Bam*HI. Contains a non-specific PCR mutated *lat* coding region. (This study)
pDA182	1.6 kb *Nco*I/*Bam*HI fragment from pDA172 inserted into pDA181 *Nco*I/*Bam*HI. Contains the wildtype *lat* coding region. (This study)
pMAL-LAT	1.6 kb *Bam*HI/*Hin*dIII fragment from pDA182 inserted into pMAL-c2 *Bam*HI/*Hin*dIII. Contains the *lat* coding region in a translational fusion with MBP. (This study)
pGEX-LAT	1.4 kb *Bam*HI/*Eco*RI fragment from pDA181 inserted into pGEX-2T *Bam*HI/*Eco*RI. Contains the mutated *lat* coding region in a translational fusion with GST. (This study)
Construction of promoter probe vectors	
pDA300	1.4 kb *Bgl*II fragment from pIJ4083 inserted into pIJ2925 *Bgl*II. Contains the pIJ4083 MCS and the *xylE* reporter gene followed by the λ t_o transcription terminator. (This study)
pDA302	1.0 kb *Pst*I fragment from pIJ4083 inserted into pKS(+) *Pst*I. Contains the fd transcription terminator. (This study)
pDA303	1.0 kb *Bgl*II/*Pst*I fragment from pDA302 inserted into pUC8 *Bam*HI/*Pst*I. Contains the fd transcription terminator. (This study)
pDA304	0.5 kb internal deletion of pDA303 by self-ligation of blunt-ended *Sph*I/*Pst*I ends. Contains the fd transcription terminator. (This study)
pDA305	0.5 kb *Eco*RI/*Hin*dIII fragment from pDA304 inserted into pDA300 *Eco*RI/*Hin*dIII. Contains the fd terminator preceding the pIJ4083 MCS and the *xylE* reporter gene followed by the λ t_o transcription terminator. (This study)
pDA307	Destruction of the unique *Xba*I site in the MCS of pFD666 by digestion with *Xba*I. Fragment ends were made blunt with Klenow polymerase and self-ligation. (This study)

Table 3. (continued)

Plasmid	Description
Construction of promoter probe vectors	
pDA308	Internal deletion of a 2.7 kb SalI fragment containing the pJV1 ori and M13 ori from pDA307. The plasmid still retains the ColE1 ori, the MCS and the bifunctional Kan^R/Neo^R marker. (This study)
pDA309	Derivative of pDA308 in which the SalI site has been mutated by insertion of the Dyl11 adapter oligonucleotide which creates an *Bgl*II site. (This study)
pDA311	3.0 kb *Pst*I fragment from pSET152 was inserted into pIJ2925 *Pst*I. Contains the ØC31 attachment site. (This study)
pDA312	Internal deletion of MCS of pDA311 by self-ligation of blunt-ended *Eco*RI/*Xba*I sites. Contains the ØC31 attachment site. (This study)
pDA313	1.9 kb *Bgl*II/*Hin*dIII fragment from pDA305 was inserted into pDA309 *Bam*HI/*Hin*dIII. (This study)
pDA320	2.4 kb *Eco*RI/*Xba*I fragment from pDA509 inserted into pDA300 *Eco*RI/*Xba*I. Contains the *tsr* resistance marker in orientation A and the FKMT terminator inserted into the *xylE* reporter gene cassette. (This study)
pDA321	2.4 kb *Eco*RI/*Xba*I fragment from pDA510 inserted into pDA300 *Eco*RI/*Xba*I. Contains the *tsr* resistance marker in orientation B and the FKMT terminator inserted into the *xylE* reporter gene cassette. (This study)
Construction of targeting vectors for gene replacement experiments	
pDA500	High copy number *Streptomyces* promoter probe plasmid pIJ4083 inserted as a 7.6 kb *Eco*RI/*Bam*HI fragment into pTZ18R *Eco*RI/*Bam*HI. Tsr^R. Amp^R. (This study)
pDA501	High copy number *Streptomyces* promoter probe plasmid pIJ486 inserted as a 6.2 kb *Eco*RI/*Bam*HI fragment into pTZ18R *Eco*RI/*Bam*HI. Tsr^R. Amp^R. (This study)
pDA502	High copy number *Streptomyces* plasmid pJOE829 inserted as a 4.7 kb *Hin*dIII fragment into pTZ18R *Hin*dIII. Bifunctional Hyg^R. Amp^R. (This study)
pDA504	1.1 kb *Bcl*I fragment from pDA501 inserted into pSL1180 *Bcl*I. Contains the *tsr* resistance marker in orientation A. (This study)
pDA505	1.1 kb *Bcl*I fragment from pDA501 inserted into pSL1180 *Bcl*I. Contains the *tsr* resistance marker in orientation B. (This study)

Table 3. (continued)

Plasmid	Description
Construction of targeting vectors for gene replacement experiments	
pDA508	3.4 kb *SacI/Bgl*II fragment from pHM8a inserted into pSL1180 *SacI/Bgl*II. Hyg^{R}. Amp^{R}. Contains the hygromycin resistance marker, *ermE** promoter and designed *Streptomyces* S/D and the FKMT terminator. (This study)
pDA509	1.3kb *Bam*HI/*Sma*I fragment from pDA508 inserted into pDA504 *Bgl*II/*Eco*ICRI. Contains the *tsr* resistance marker in orientation A followed by the FKMT terminator. (This study)
pDA510	1.3kb *Bam*HI/*Sma*I fragment from pDA508 inserted into pDA505 *Bgl*II/*Eco*ICRI. Contains the *tsr* resistance marker in orientation B followed by the FKMT terminator. (This study)
pDA511	2.4 kb *Eco*RI/*Xba*I fragment from pDA509 inserted into pDA300 *Eco*RI/*Xba*I. Contains the *tsr* resistance marker in orientation A and the FKMT terminator inserted into the *xylE* reporter gene cassette. (This study)
pDA512	2.4 kb *Eco*RI/*Xba*I fragment from pDA510 inserted into pDA300 *Eco*RI/*Xba*I. Contains the *tsr* resistance marker in orientation B and the FKMT terminator inserted into the *xylE* reporter gene cassette. (This study)
pDA513	0.3 kb *Eco*ICRI/*Nco*I fragment from pDA508 inserted into pSKN *Sma*I/*Nco*I. Contains the expression cassette with *ermE** promoter and designed S/D. (This study)
pDA515	1.5 kb *Eco*RI/*Pst*I fragment from pOW309 inserted into pUC119 *Eco*RI/*Pst*I. Bifunctional Apr^{R}. Amp^{R}. Contains the *apr* resistance marker. (This study)
pDA516	Derivative of pDA515 in which the Dyl19 adapter oligonucleotide was inserted at the unique *Sph*I site within the MCS creating a new *Spe*I site. Bifunctional Apr^{R}. Amp^{R}. Contains the *apr* resistance marker preceded by amber stop codons in all three reading frames (*apr* cassette). (This study)
pDA517	1.5 kb *Eco*RI/*Hin*dIII fragment from pDA516 inserted into pKS(+) *Eco*RI/*Hin*dIII. Bifunctional Apr^{R}. Amp^{R}. Contains the *apr* cassette. (This study)
pDA520	Blunt-ended 1.5 kb *Hin*dIII/*Sma*I fragment from pDA517 inserted into pDA134 linearized at the unique *Eco*47III site. Contains the *apr* cassette inserted into the *ccaR* gene in orientation B. (This study)
pDA521	Blunt-ended 1.5 kb *Hin*dIII/*Sma*I fragment from pDA517 inserted into pDA134 linearized at the unique *Eco*ICRI site. Contains the *apr* cassette inserted into the *ccaR* gene in orientation A. (This study)

Table 3. (continued)

Plasmid	Description
Construction of targeting vectors for gene replacement experiments	
pDA522	Blunt-ended 1.5 kb *Hin*dIII/*Sma*I fragment from pDA517 inserted into pDA134 linearized at the unique *Eco*ICRI site. Contains the *apr* cassette inserted into the *ccaR* gene in orientation B. (This study)
pDA523	Blunt-ended 1.5 kb *Hin*dIII/*Sma*I fragment from pDA517 inserted into pDA134 linearized at the unique *Nco*I site (blunt-ended with Klenow polymerase). Contains the *apr* cassette inserted into the *orf11* gene in orientation A. (This study)
pDA524	Blunt-ended 1.5 kb *Hin*dIII/*Sma*I fragment from pDA517 inserted into pDA134 linearized at the unique *Nco*I site (blunt-ended with Klenow polymerase). Contains the *apr* cassette inserted into the *orf11* gene in orientation B. (This study)
pDA525	Blunt-ended 1.5 kb *Hin*dIII/*Sma*I fragment from pDA517 inserted into pDA134 linearized at the unique *Not*I site (blunt-ended with Klenow polymerase). Contains the *apr* cassette inserted into the *blp* gene in orientation A. (This study)
pDA526	Blunt-ended 1.5 kb *Hin*dIII/*Sma*I fragment from pDA517 inserted into pDA134 linearized at the unique *Not*I site (blunt-ended with Klenow polymerase). Contains the *apr* cassette inserted into the *blp* gene in orientation B. (This study)
pDA527	2.3 kb *Nru*I/*Eco*RI fragment from pDA120 inserted into pDA504 digested with *Bam*HI/ *Eco*RI (the *Bam*HI end was blunt-ended with Klenow polymerase). Contains the *tsr* marker in orientation A flanked by the *orf11* and *blp* genes. (This study)
pDA528	2.3 kb *Nru*I/*Eco*RI fragment from pDA120 inserted into pDA505 digested with *Bam*HI/ *Eco*RI (the *Bam*HI end was blunt-ended with Klenow polymerase). Contains the *tsr* marker in orientation B flanked by the *orf11* and *blp* genes. (This study)
pDA529	2.8 kb *Bam*HI/*Xba*I fragment from pDA130 inserted into pDA527 *Bgl*II/*Xba*I. Contains the *tsr* marker in orientation A flanked by the *orf11* and *blp* genes on one side and *cmcH* and part of the *cefF* gene on the other. (This study)
pDA530	2.8 kb *Bam*HI/*Xba*I fragment from pDA130 inserted into pDA528 *Bgl*II/*Xba*I. Contains the *tsr* marker in orientation B flanked by the *orf11* and *blp* genes on one side and *cmcH* and part of the *cefF* gene on the other. (This study)
pDA531	1.7 kb *Eco*RI/*Apa*I fragment (the *Eco*RI end was blunt-ended with Klenow polymerase) from pDA172 was inserted into pDA504 digested with *Bam*HI/*Apa*I (the *Bam*HI end was blunt-ended with Klenow polymerase). Contains the *tsr* marker in orientation A flanked by the *lat* gene. (This study)

Table 3. (continued)

Plasmid	Description
Construction of targeting vectors for gene replacement experiments	
pDA532	1.7 kb *Eco*RI/*Apa*I fragment (the *Eco*RI end was blunt-ended with Klenow polymerase) from pDA172 was inserted into pDA505 digested with *Bam*HI/*Apa*I (the *Bam*HI end was blunt-ended with Klenow polymerase). Contains the *tsr* marker in orientation B flanked by the *lat* gene. (This study)
pDA533	2.8 kb *Bam*HI/*Xba*I fragment from pDA130 inserted into pDA531 *Bgl*II/*Xba*I. Contains the *tsr* marker in orientation A flanked by the *lat* gene on one side and *cmcH* and part of the *cefF* gene on the other. (This study)
pDA534	2.8 kb *Bam*HI/*Xba*I fragment from pDA130 inserted into pDA532 *Bgl*II/*Xba*I. Contains the *tsr* marker in orientation B flanked by the *lat* gene on one side and *cmcH* and part of the *cefF* gene on the other. (This study)
pDA535	1.0 kb *tsr* marker *Nco*I/*Eco*47III fragment from pDA504 was inserted into pDA170 *Nco*I/*Eco*47III. Contains the *tsr* marker in orientation A flanked by the rest of the *lat* gene on one side and *lat* promoter and the *blp* gene and part of the *orf11* gene on the other. (This study)
pDA536	1.0 kb *tsr* marker *Nco*I/*Eco*47III fragment from pDA505 was inserted into pDA170 *Nco*I/*Eco*47III. Contains the *tsr* marker in orientation A flanked by the rest of the *lat* gene on one side and *lat* promoter and the *blp* gene and part of the *orf11* gene on the other. (This study)
pDA537	Destruction of the MCS *Eco*RI site by self-ligation of the blunt-ended *Eco*RV/*Bam*HI sites in the MCS of pDA171. Contains the *blp* and *lat* genes and part of the *orf11* gene. (This study)
pDA538	2.3 kb *tsr Eco*RI/*Hin*dIII (the *Hin*dIII end was blunt-ended with Klenow polymerase) of pDA509 was inserted into pDA537 *Eco*RI/*Eco*47III. Contains the *tsr* marker in orientation A and the FKMT terminator flanked by the rest of the *lat* gene on one side and the *blp* gene and part of the *orf11* gene on the other. (This study)
pDA539	2.3 kb *tsr Eco*RI/*Hin*dIII (the *Hin*dIII end was blunt-ended with Klenow polymerase) of pDA510 was inserted into pDA537 *Eco*RI/*Eco*47III. Contains the *tsr* marker in orientation B and the FKMT terminator flanked by the rest of the *lat* gene on one side and the *blp* gene and part of the *orf11* gene on the other. (This study)
pDA540	Destruction of the unique *Kpn*I/*Acc*65I site by digestion of pDA173 with *Acc*65I, fragment ends were made blunt with Klenow polymerase and self-ligated. Contains a mutated *lat* gene altered at the *Kpn*I site. (This study)

Table 3. (continued)

Plasmid	Description
Construction of targeting vectors for gene replacement experiments	
pDA541	Deletion of 354 bp by digestion of pDA173 with *Acc*65I and *Eco*47III, fragment ends were made blunt with Klenow polymerase and self-ligated. Contains a mutant *lat* gene with an internal deletion. (This study)
pDA542	1.6 kb *Eco*RI/*Xba*I fragment from pDA540 inserted into pDA170 *Eco*RI/*Xba*I. Contains a mutated *lat* gene altered at the *Kpn*I site as well as the *blp* gene and part of the *orf11* gene. (This study)
pDA543	1.2 kb *Eco*RI/*Xba*I fragment from pDA541 inserted into pDA170 *Eco*RI/*Xba*I. Contains a mutated *lat* gene with an internal deletion as well as the *blp* gene and part of the *orf11* gene. (This study)
pDA551	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA521 *Hin*dIII. The *E. coli*/*Streptomyces* shuttle plasmid contains the *orf11* and *blp* genes and the *apr* cassette inserted into the unique *Eco*ICRI site of the *ccaR* gene in orientation A. (This study)
pDA552	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA522 *Hin*dIII. The *E. coli*/*Streptomyces* shuttle plasmid contains the *orf11* and *blp* genes and the *apr* cassette inserted into the unique *Eco*ICRI site of the *ccaR* gene in orientation B. (This study)
pDA553	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA523 *Hin*dIII. The *E. coli*/*Streptomyces* shuttle plasmid contains the *ccaR* and *blp* genes and the *apr* cassette inserted into the unique *Nco*I site of the *orf11* gene in orientation A. (This study)
pDA554	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA524 *Hin*dIII. The *E. coli*/*Streptomyces* shuttle plasmid contains the *ccaR* and *blp* genes and the *apr* cassette inserted into the unique *Nco*I site of the *orf11* gene in orientation B. (This study)
pDA555	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA525 *Hin*dIII. The *E. coli*/*Streptomyces* shuttle plasmid contains the *ccaR* and *orf11* genes and the *apr* cassette inserted into the unique *Not*I site of the *blp* gene in orientation A. (This study)
pDA556	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA526 *Hin*dIII. The *E. coli*/*Streptomyces* shuttle plasmid contains the *ccaR* and *orf11* genes and the *apr* cassette inserted into the unique *Not*I site of the *blp* gene in orientation B. (This study)
pDA557	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA529 *Hin*dIII. The *E. coli*/*Streptomyces* shuttle plasmid contains the *tsr* marker, in orientation A, flanked by the *orf11* and *blp* genes on one side and the *cmcH* and part of the *cefF* genes on the other. (This study)

Table 3. (continued)

Plasmid	Description
Construction of targeting vectors for gene replacement experiments	
pDA558	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA530 *Hin*dIII. The *E. coli/Streptomyces* shuttle plasmid contains the *tsr* marker, in orientation B, flanked by the *orf11* and *blp* genes on one side and the *cmcH* and part of the *cefF* genes on the other. (This study)
pDA559	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA533 *Hin*dIII. The *E. coli/Streptomyces* shuttle plasmid contains the *tsr* marker, in orientation A, flanked by the *lat* gene on one side and the *cmcH* and part of the *cefF* genes on the other. (This study)
pDA560	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA534 *Hin*dIII. The *E. coli/Streptomyces* shuttle plasmid contains the *tsr* marker, in orientation B, flanked by the *lat* gene on one side and the *cmcH* and part of the *cefF* genes on the other. (This study)
pDA561	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA535 *Hin*dIII. The *E. coli/Streptomyces* shuttle plasmid contains the *tsr* marker, in orientation A, flanked by part of the *lat* gene on one side and the *blp* gene and parts of the *lat* and *orf11* genes on the other. (This study)
pDA562	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA536 *Hin*dIII. The *E. coli/Streptomyces* shuttle plasmid contains the *tsr* marker, in orientation B, flanked by part of the *lat* gene on one side and the *blp* gene and parts of the *lat* and *orf11* genes on the other. (This study)
pDA563	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA538 *Hin*dIII. The *E. coli/Streptomyces* shuttle plasmid contains the *tsr* marker in orientation A followed by the FKMT terminator, flanked by part of the *lat* gene on one side and the *blp* gene and parts of the *lat* and *orf11* genes on the other. (This study)
pDA564	4.7 kb *Hin*dIII fragment from pDA502 inserted into pDA539 *Hin*dIII. The *E. coli/Streptomyces* shuttle plasmid contains the *tsr* marker in orientation B followed by the FKMT terminator, flanked by part of the *lat* gene on one side and the *blp* gene and parts of the *lat* and *orf11* genes on the other. (This study)
pDA565	6.2 kb *Bgl*II fragment from pDA501 inserted into pDA542 *Bam*HI. The *E. coli/Streptomyces* shuttle plasmid contains the *lat* gene with a frameshift mutation at the central unique *Kpn*I site. (This study)
pDA566	6.2 kb *Bgl*II fragment from pDA501 inserted into pDA543 *Bam*HI. The *E. coli/Streptomyces* shuttle plasmid contains the *lat* gene with a deletion mutation from the *Eco*47III site to the *Kpn*I site. (This study)

Table 3. (continued)

Plasmid	Description
Cloning of cephamycin cluster fragments into pUWL-KS	
pDA601	0.2 kb *Eco*RI/*Xba*I fragment of pDA175 was inserted into pUWL-KS *Eco*RI/*Xba*I. Contains the *lat* promoter. (This study)
pDA602	1.8 kb *Eco*RI/*Bam*HI fragment of pDA172 was inserted into pUWL-KS *Eco*RI/*Bam*HI. Contains the *lat* gene. (This study)
Complementation vectors using pSET152	
pDA1000	1.1 kb *Eco*RI/*Xba*I fragment from pDA504 inserted into pSET152 *Eco*RI/*Xba*I. Contains the *tsr* resistance marker. (This study)
pDA1001	3.7 kb *Eco*RI/*Bam*HI fragment from pDA120 inserted into pSET152 *Eco*RI/*Bam*HI. Contains the wildtype *ccaR*, *orf11* and *blp* genes. (This study)
pDA1002	3.7 kb *Eco*RI/*Bam*HI fragment from pDA140 inserted into pSET152 *Eco*RI/*Bam*HI. Contains a mutant *ccaR* gene and wildtype *orf11* and *blp* genes inserted into pSET152. (This study)
pDA1003	3.7 kb *Eco*RI/*Bam*HI fragment from pDA141 inserted into pSET152 *Eco*RI/*Bam*HI. Contains a mutant *orf11* gene and wildtype *ccaR* and *blp* genes inserted into pSET152. (This study)
pDA1004	3.7 kb *Eco*RI/*Bam*HI fragment from pDA142 inserted into pSET152 *Eco*RI/*Bam*HI. Contains a mutant *blp* gene and wildtype *ccaR* and *orf11* genes inserted into pSET152. (This study)
pDA1005	1.7 kb *Eco*RI/*Xba*I fragment from pDA150 inserted into pSET152 *Eco*RI/*Xba*I. Contains the *ccaR* gene and part of the *orf11* gene. (This study)
pDA1006	1.4 kb *Eco*RI/*Nru*I fragment from pDA150 inserted into pSET152 *Eco*RI/*Eco*RV. Contains the *ccaR* gene. (This study)
pDA1007	1.4 kb *Bam*HI/*Nru*I fragment from pDA143 inserted into pSET152 *Bam*HI/*Eco*RV. Contains a mutant *ccaR* gene altered at the *Eco*47III site. (This study)
pDA1008	1.4 kb *Bam*HI/*Nru*I fragment from pDA140 inserted into pSET152 *Bam*HI/*Eco*RV. Contains a mutant *ccaR* gene altered at the *Eco*ICRI site. (This study)
pDA1010	1.0 kb *Bam*HI/*Eco*ICRI fragment from pDA160 inserted into pSET152 *Bam*HI/*Eco*RV. Contains a mutant *ccaR* gene deleted in the C-terminus at the *Eco*ICRI. (This study)

Table 3. (continued)

Plasmid	Description
Complementation vectors using pSET152	
pDA1011	1.0 kb *Bam*HI/*Stu*I fragment from pDA161 inserted into pSET152 *Bam*HI/*Eco*RV. Contains a mutant *ccaR* gene deleted in the C-terminus at the *Stu*I site. (This study)
pDA1012	1.1 kb *Eco*RI/*Xba*I fragment from pDA163 inserted into pSET152 *Eco*RI/*Xba*I. Contains a mutant *ccaR* gene deleted in the C-terminus between the *Stu*I and *Kpn*I sites. (This study)
pDA1013	1.1 kb *Eco*RI/*Xba*I fragment from pDA164 inserted into pSET152 *Eco*RI/*Xba*I. Contains a mutant *ccaR* gene deleted in the C-terminus between the *Stu*I and *Kpn*I sites. (This study)
pDA1014	1.1 kb *Bam*HI/*Nru*I fragment from pDA162 inserted into pSET152 *Bam*HI/*Eco*RV. Contains a mutant *ccaR* gene deleted in the C-terminus at the *Kpn*I. (This study)
pDA1020	1.4 kb *Bam*HI/*Hin*dIII (the *Hin*dIII end was blunt-ended with Klenow polymerase) fragment from pDA153 inserted into pSET152 *Bam*HI/*Eco*RV. Contains the *ccaR* gene of *S. clavuligerus dclX*68. (This study)
pDA1021	1.4 kb *Bam*HI/*Hin*dIII (the *Hin*dIII end was blunt-ended with Klenow polymerase) fragment from pDA154 inserted into pSET152 *Bam*HI/*Eco*RV. Contains the *ccaR* gene of *S. clavuligerus* NTG 26. (This study)
pDA1022	The 1.4 kb *Bam*HI/*Hin*dIII (the *Hin*dIII end was blunt-ended with Klenow polymerase)fragment from pDA155 inserted into pSET152 *Bam*HI/*Eco*RV. Contains the *ccaR* gene of *S. clavuligerus* NTG 32. (This study)
pDA1023	1.4 kb *Bam*HI/*Hin*dIII (the *Hin*dIII end was blunt-ended with Klenow polymerase)fragment from pDA156 inserted into pSET152 *Bam*HI/*Eco*RV. Contains the *ccaR* gene of *S. clavuligerus* NTG 83. (This study)
pDA1030	2.0 kb *Eco*RI/*Xba*I fragment from pDA151 inserted into pSET152 *Eco*RI/*Xba*I. Contains the *ccaR* gene and also part of the *orf11* gene with the *ermE** promoter placed in tandem with the *ccaR* promoter. (This study)
pDA1050	1.8 kb *Eco*RI/*Bam*HI fragment from pDA172 inserted into pSET152 *Eco*RI/*Bam*HI. Contains the *lat* gene. (This study)
pDA1053	2.0 kb *Eco*RI/*Xba*I fragment from pDA177 inserted into pSET152 *Eco*RI/*Xba*I. Contains the *lat* coding region fused to the *ermE** promoter and S/D. (This study)

Table 3. (continued)

Plasmid	Description
***xylE* promoter probe constructs**	
pDA2000	3.0 kb *Bgl*II fragment from pDA312 was inserted into pDA313. Creates an *E. coli/Streptomyces* shuttle plasmid functioning as a single-copy promoter probe vector in *Streptomyces*. High copy number cloning vector with small MCS in *E. coli*. Bifunctional Kan^R/Neo^R. (This study)
pDA2001	225 bp *Eco*RI/*Xba*I fragment from pDA175 inserted into pDA2000 *Eco*RI/*Xba*I. Contains the *lat* promoter. (This study)
pDA2002	1.1 kb *Eco*RI/*Kpn*I fragment from pDA174 inserted into pDA2000 *Eco*RI/*Kpn*I. Contains the *lat* promoter and some of the *lat* coding region. (This study)
pDA2003	1.6 kb *Eco*RI/*Xba*I fragment from pDA173 inserted into pDA2000 *Eco*RI/*Xba*I. Contains the entire *lat* gene inserted into pDA2000. (This study)
pDA2004	1.8 kb *Eco*RI/*Bam*HI fragment from pDA172 inserted into pDA2000 *Eco*RI/*Bam*HI. Contains the entire *lat* gene and the downstream terminator. (This study)
pDA2007	3.9 kb *Eco*RI/*Bgl*II fragment from pDA320 cloned into pSET152 *Eco*RI/*Bam*HI. Contains the *tsr* resistance marker in orientation A and the FKMT terminator inserted into the *xylE* reporter gene cassette. (This study)
pDA2008	3.9 kb *Eco*RI/*Bgl*II fragment from pDA321 cloned into pSET152 *Eco*RI/*Bam*HI. Contains the *tsr* resistance marker in orientation B and the FKMT terminator inserted into the *xylE* reporter gene cassette. (This study)
pDA2500	1.9 kb *Bgl*II/*Hin*dIII fragment from pDA305 was inserted into pDA307 *Bam*HI/*Hin*dIII. Creates an *E. coli/Streptomyces* shuttle plasmid functioning as a mid-copy number promoter probe vector in *Streptomyces*. High copy number cloning vector with small MCS in *E. coli*. Bifunctional Kan^R/Neo^R. (This study)

II.3.5.3 Preparation of antigen for injection into rabbits

The purified MBP-fusion proteins as well as the purified IPNS protein (65), approximately 150 µg of each, were run on separate 7.5% SDS-PAG or 10% SDS-PAG, respectively. The gels were stained for 30 minutes in 0.1% (w/v) Coomassie blue in milliQ water before destaining in water for 10 minutes. The stained protein bands were cut out from the gel using a razor blade with care taken to minimize the amount of extra acrylamide. The protein/acrylamide fragments were homogenized in phosphate buffered saline [80 mM Na_2HPO_4, 20 mM NaH_2PO_4 and 100 mM NaCl (194)] by rapid transfer between two syringes using a two-way luer-lock connector. This slurry was mixed with an equal volume of Freund's complete adjuvant for the first injection or Freund's incomplete adjuvant for the subsequent injections (92). Immunization procedures were done by the Department of Biological Sciences Animal Services group in accordance with the animal welfare guidelines of the University of Alberta.

II.3.5.4 Collection of polyclonal antibody containing sera

The rabbit blood was allowed to clot in the vacutainer harvest tubes before transferring to sterile centrifuge tubes. The blood was centrifuged for 10 minutes before transferring the serum to a fresh sterile flask. The serum was aliquoted into sterile tubes and stored at -70°C. A working solution of about 1 ml was stored at 4°C.

II.3.5.5 Purification of antibodies from western blots

The antisera raised versus the MBP-fusion proteins showed some cross-reactivity with other proteins in the cell-extracts. Therefore attempts were made to purify the antibodies by immunoaffinity (92). The purification of antibodies specific for either the CcaR or LAT proteins but not the MBP part of the fusion was attempted using CcaR and LAT fusion proteins generated with the pGEX-2T expression vector which fused LAT or CcaR to the glutathione-S-transferase (GST) protein. The artificial *Bam*HI sites generated on the 5' end of each gene just before the start codon created a translational fusion with the GST protein when inserted into pGEX-2T.

II.3.5.5.1 Production and purification of GST-fusion proteins

Overnight cultures of pGEX-CcaR and pGEX-LAT were grown in 2x YT before diluting 1:100 into 50 ml of prewarmed 2x YT and incubating for 1 hour at 37°C. Cultures were induced by adding IPTG to a final concentration of 0.5 mM and then incubated at 37°C for 4 hours.

Both of the GST-fusion proteins were produced as inclusion bodies after induction with IPTG. Cells containing the GST-fusion protein inclusion bodies were harvested and broken in STE but the inclusion bodies were pelleted during the centrifugation. Repeated sonication, centrifugation, and washing with STE and STE + 1.0 % Tween 20 (three washes each) created a preparation of inclusion bodies that was quite pure as judged by SDS-PAGE.

II.3.5.5.2 Immunoaffinity purification of antibodies

Each of the inclusion body preparations was mixed with sample buffer and electrophoresed on 4 separate 10% SDS-PAGs and transferred to PVDF membranes (see II.2.4.3.2). The transferred PVDF membranes were blocked in 5% (w/v) skim milk in TBST for one hour at room temperature and then washed by gentle agitation with TBSTM: two quick rinses, one 15 minute wash followed by two 5 minute washes. The primary antibody was added to fresh TBSTM at a 1:50 dilution and incubated for 60 minutes. The blots were again washed with TBSTM: two quick rinses, one 15 minute wash followed by two 5 minute washes. Vertical strips were removed from each end, incubated with secondary antibody, developed and exposed to film (see II.2.4.3.3). The remaining part of the membrane was lined up with the developed X-ray film, portions of the membrane containing the antigen and primary antibody complex were removed with a razor blade and cut into small pieces. These pieces were incubated with 1 ml of 100 mM glycine pH 2.5 for 10 minutes before adding one-tenth volume of 1 M Tris-HCl pH 8.0 to neutralize. The tubes were pooled and the solution containing the purified antibodies were filter sterilized and stored at 4^{0}C. This procedure was only successful for the purification of LAT-specific antibodies.

II.4 Recombinant DNA Techniques

The following are adaptations/combinations of the protocols outlined in Sambrook et al. (194), Hopwood et al. (97) and the Promega protocols and applications guide (184). These protocols were used routinely throughout the work.

A list of the oligonucleotides used as adapters, PCR primers, or as custom sequencing primers is provided in Table 4.

II.4.1 Organic solvent extraction of DNA and ethanol precipitation

The restriction enzyme(s) used to digest DNA that was to be used for cloning were removed by phenol extraction. The total volume of the restriction digest was taken up to 200 µl with TE and 200 µl of a phenol/chloroform/isoamyl alcohol mixture was

Table 4. Oligonucleotides list

Name	Sequence	Description
Apa8	GATCTCCATGGA	*Bam*HI-*Nco*I adapter which creates a *Nco*I site
Dyl9	GCCGGATCCATGGGCGAAGCAGCACGCCACCC	5' *lat* primer which generates an upstream *Bam*HI site
Dyl10	GCCGAATTCGCGTCAGACGCTCTCGGCGACCGG	3' *lat* primer which generates a downstream *Eco*RI site
Dyl11	TCGATAGATCTA	*Sal*I-*Bgl*II adapter which creates a *Bgl*II site
Dyl12	GCCGAATTCGCGGTTTCAGGCCGGGGTACGAC	3' *dclX* primer which generates a downstream *Eco*RI site
Dyl13	GCCGGATCCATCAACACCTGGAATGATGTGACG	5' *dclX* primer which generates an upstream *Bam*HI site
Dyl14	CGCACATACTCGCCGCCCTTC	Custom primer for pDA103 sequencing
Dyl17	CCGCGAAGAAGGGCGGACACC	Custom primer for pDA102 sequencing
Dyl19	CTAGACTAGTCTAGCATG	*Sph*I-*Spe*I adapter which contains amber stop codons
Dyl21	GTCAGGTTTGTTTCCGCCAACGTGGCG	pDA2000 extra-long reverse MCS PCR primer
Dyl22	GGGCGTGGACTTCTCCCAG	Custom primer for pOW309 Δ*Kpn*I sequencing
Dyl23	GGCCCGCTGGGAATCCGTC	Custom primer for pDA102 sequencing
Dyl24	CCTTCTGACCGGACAGGCCG	Custom primer for pOW309 Δ*Kpn*I sequencing
Dyl28	AGGCGAGGGCTTCCGGCGCC	Custom primer for pOW309 Δ*Kpn*I sequencing
Dyl31	CGCCAGGGTTTTCCCAGTCACGAC	pSET152 extra-long forward MCS PCR primer
Dyl32	GAGCGGATAACAATTTCACACAGGA	pSET152 extra-long reverse MCS PCR primer
Dyl33	ACTAACGTTTGGAAAGACGACAAAAC	pDA2000 extra-long forward MCS PCR primer
Wsc13	TGGCCATGGCGTACACGGGCTTCACGCCCG	5' *blp* primer

added and vigorously mixed. After a 2 minute microcentrifugation the upper phase was transferred to a tube containing 200 µl of a mixture containing chloroform/isoamyl alcohol and again mixed vigorously. After another 2 minute microcentrifugation 180 µl of the upper phase was added to 0.5 µl of glycogen, 90 µl of 7.5 M NH_4OAc and 360 µl of 98% ethanol. If the digested DNA was going to be dephosphorylated the 180 µl of the upper phase was added to 0.5 µl of glycogen, 18 µl of 3 M NaOAc pH 7.0 and 360 µl of 98% ethanol. The tubes were incubated at -70ºC for at least 20 minutes followed by a 10 minute microcentrifugation. The supernatant was removed, replaced with 200 µl of 70% ethanol and washed for five minutes. The tubes were microcentrifuged for 1 minute and the ethanol was removed before the DNA was vacuum dried.

II.4.2 Gel purification of DNA fragments

Purification of DNA fragments from agarose gel blocks was done using the GlassMax DNA isolation system. The agarose blocks were resuspended in 500 to 1000 µl of binding solution and incubated at 50ºC until the agarose had dissolved. Glass resin, 10 to 20 µl per tube depending on the amount of DNA, was added to each tube and mixed slowly by inversion on a tube roller at room temperature for 1 to 2 hours. The resin was precipitated by a 20 second microcentrifugation and washed 3 times with 500 µl of wash solution. After the final wash all of the wash solution was removed from the tube. The pellet was resuspended with a volume of TE two times the volume of glass resin used and heated at 50ºC for 3 minutes. The total volume was increased to 200 µl with TE and microcentrifuged for 2 minutes before 180 µl was removed and ethanol precipitated.

II.4.3 Dephosphorylation of DNA ends

Cloning vectors linearized by digestion with a single restriction enzyme were routinely treated with alkaline phosphatase to dephosphorylate the ends and often vectors that were linearized with two different restriction enzymes were also dephosphorylated. This was done to decrease the amount of vector religation.

In the dephosphorylation procedure the dried DNA pellet, up to 5 µg of DNA, was resuspended in 3 µl of 10X alkaline phosphatase buffer (supplied with enzyme) and 25.5 µl of water before the reaction was initiated by adding 1 µl of calf intestinal alkaline phosphatase (1 unit/µl) and incubated at 37ºC. After a 15 minute incubation another 1 µl of enzyme was added and the tube was again incubated for 15 minutes. The dephosphorylation reaction was terminated by heating to 65ºC for 15 minutes and then the reaction mixture was organic solvent extracted and ethanol precipitated.

II.4.4 Generation of blunt-ends with Klenow or T4 DNA polymerase

Often it was necessary to convert a cohesive end to a blunt-end on a DNA fragment for cloning if a compatible unique cohesive site was not available. This procedure was done before the insert DNA was gel purified.

The dried DNA pellet, up to 5 µg of DNA, was resuspended in 3 µl of 10X Klenow buffer (0.5 M Tris-HCl pH 7.6 and 0.1 M $MgCl_2$) and 25.5 µl of water before the reaction was initiated by adding 0.5 µl of the Klenow fragment (2 units/µl) and incubated at 37°C (for 3' overhangs only). After 10 minutes 1 µl of 0.25 mM dNTPs was added and incubated for 15 minutes (both 3' and 5' overhangs). The reaction was terminated by organic solvent extraction and ethanol precipitation.

The *Pst*I site proved to be a difficult site to "chew-back" and "filled-in" so the site was made blunt-ended with T4 DNA polymerase. The dried DNA pellet was resuspended in 3 µl of 10X restriction buffer H (from Boehringer Mannheim) and 26.5 µl of water before the reaction was initiated by the addition of 0.5 µl of T4 DNA polymerase (1unit/µl) and incubated at 37°C for 15 minutes.

II.4.5 Random primer labeling of hybridization probes

A gel purified fragment of DNA, 0.5 to 1 µg of DNA in up to a total volume of 8 µl, was heated at 95°C for 2 minutes and then cooled on ice immediately. Added to the denatured DNA were 2 µl of random hexanucleotide primer (50 OD/ml), 4 µl of 5X buffer [25 mM MgCl2, 10 mM DTT, 2 mg/ml BSA, 1 M HEPES (pH 6.6) and 250 mM Tris-HCl (pH 8) (71, 72)], 1 µl of each 1 mM stock of cold dNTP and 1 µl of [α-^{32}P] labeled dNTP (usually dCTP) and water to a total volume of 20 µl. This mixture was heated to 90°C for 2 minutes and then incubated at 37°C for 15 minutes before 1 µl of Klenow enzyme was added and the mixture was allowed to continue incubating at 37°C for another 2 hours. The reaction was stopped by the addition of 100 µl TE plus 0.1% (w/v) SDS. Since incorporation of label by this method was very good the probes were not purified before being used in hybridization procedures.

II.4.6 Ligation of DNA Molecules

Ligation reactions were set up such that the insert, usually gel purified, was present in a 3 to 1 molar ratio to the vector. The vector DNA was often dephosphorylated using alkaline phosphatase and in many cases completed ligation reaction mixtures were digested with restriction enzymes unique to the MCS fragment being removed before being transformed.

Ligation reactions involving large fragments of DNA (insert plus vector greater than 12 kb) with compatible cohesive ends were done using 10X Boehringer Mannheim ligation buffer and incubated at 16°C for at least 6 hours. Cohesive/blunt end ligation reactions (one end of the fragment is cohesive and the other end is blunt) or cohesive end ligation reactions of smaller pieces of DNA were done using 5X BRL homemade ligation buffer (250 mM Tris-HCl pH 7.6, 50 mM $MgCl_2$, 25% (w/v) PEG 8000, 5 mM ATP and 5 mM DTT) and incubated at 16°C for at least 6 hours. These ligation reactions were heated at 55°C for five minutes and then cooled on ice for 5 minutes before the ligase buffer and ligase were added.

Blunt end ligation reactions were done using the 5X BRL ligation buffer and incubated at 22°C for at least 6 hours. For self-ligation reactions, blunt or cohesive, the ligation reactions were done in 10X ligation buffer and incubated at 22°C for 1 to 2 hours.

In some cases an oligonucleotide adapter was inserted into a linearized vector. All adapters were designed to contain an internal region of self-complementarity, containing a new restriction site, which after annealing gave a double stranded molecule with 5' or 3' cohesive ends. Adapter oligonucleotides were self-annealed by combining 600 pmoles of oligonucleotide with 10X annealing buffer (500 mM NaCl, 100 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$ and 10 mM DTT) and water, the mixture was heated to 65°C for 10 minutes before slow cooling to room temperature. An aliquot of this annealed adapter was ligated together with non-dephosphorylated DNA linearized at a single site. Often these adapter oligonucleotides were designed to mutate the restriction site they were inserted into and the resulting ligation mixtures could be digested with the original linearizing enzyme to decrease the background of self-ligated plasmid.

Typically, ligation mixtures were used to transform *E. coli* by the $MnCl_2$ method. Electroporation was often reserved for ligation reactions resulting in plasmids greater than 12 kb or ligation reactions that were unsuccessful when transformed into $MnCl_2$ cells.

II.4.7 Creation of exoIII deletion clones

II.4.7.1 Generation of exoIII deletions for sequencing

The plasmid template to be digested was isolated with the Wizard miniprep system to give the very high percentage of supercoiled plasmid DNA necessary for the experiment. The experimental design was outlined in the Erase-a-Base kit protocol and each experiment would differ only by the amount of template that had to be digested and therefore the number of reactions to be done. Template DNA was first digested with a restriction enzyme which would leave 3' overhang. Complete restriction enzyme digestion was confirmed by agarose gel electrophoresis before the second digestion was started. The restriction enzyme used in the second digestion would leave a 5' overhang that was digested by the exoIII. The recipes for the stock buffers included with the kit were provided by the Erase-a-Base manual (184). S1 nuclease and exoIII enzymes provided by the kit were found to be unreliable so the corresponding enzymes from Boehringer Mannheim were used in their place. In most cases, the reactions for both orientations were done in parallel.

The extracted and precipitated template was resuspended in milliQ water and an aliquot containing approximately 2.5 µg of template was added to 2.6 µl of 10X exoIII buffer and made up to 23 µl with milliQ water. This provided enough exoIII mixture for 8 time points. When incubated at 37ºC exoIII has a predicted degradation rate of 400 bp/minute, so 8 time points taking a sample every 45 seconds would be enough to remove all of a 2 kb insert. The S1 mixture was made for 20 reactions, 20 µl of 10X S1 buffer, 130 µl of water and 0.2 µl of S1 nuclease (400 units/µl). Aliquots of 7.5 µl were transferred into fresh tubes and stored on ice. If a large number of time points were being handled 15 µl was dispensed into each tube and the aliquots removed at two sequential time points were added to the same tube and all subsequent reactions were doubled as well. The exoIII tube was preheated to 37ºC, using the PCR machine as a temperature block, before 2.9 µl of exoIII (100 units/µl) was added, mixed well and returned to 37ºC. The first time point was at 65 seconds and 2.5 µl of the exoIII reaction tube was removed and added to its corresponding tube containing the S1 mixture and returned to the ice. Subsequent time samples were removed every 45 seconds until completion. The first time point was delayed 20 seconds since the exoIII reaction has a small lag before proceeding. All of the S1 tubes were then incubated at 22ºC for 30 minutes before 1 µl of S1 stop mix was added. The samples were then heated to 70ºC for 10 minutes after which the samples were microcentrifuged for 20 seconds and 2 µl was removed to run on an agarose gel before the tubes were returned to 37ºC. Klenow mixture was prepared in which 20 µl of 10X Klenow buffer was mixed with 0.6 µl of Klenow enzyme (2 units/µl) and 1 µl of this mixture was added to each S1 tube and incubated for 5 minutes. Then 1 µl of dNTPs was added to each tube and they were then incubated for 10 minutes. A ligase mixture was then made, which contained 80 µl of 10X ligase buffer, 80 µl of 50% PEG, 8 µl of 0.1M DTT, 632 µl of water and 1.5 µl of ligase (1 unit/µl). After the Klenow incubation was complete, 40 µl of ligase mixture was added to each tube and the ligation reaction was allowed to proceed at least 4 hours at 22ºC before transforming one quarter of the reaction into *E. coli* $MnCl_2$ competent cells.

In a successful exoIII deletion experiment the agarose gel run after the S1 digestion step would show a series of stepwise decreases in the size of the template molecule.

II.4.7.1.1 Screening deletion clones for sequencing

After the ligation mixtures from the exoIII digestion procedure were transformed into competent cells, six clones from each transformation plate were used to isolate DNA. These clones, the original template and pKS(+) or pTZ18R vector without an insert were digested with *Pvu*II. *Pvu*II sites flank the MCS of the pKS(+) or pTZ18R vectors so *Pvu*II digestion should release the insert and give an estimate of the size of each deletion clone, provided that the insert does not have any *Pvu*II sites and that insert size is not excessively large (i.e. >5 kb). Sequential clones estimated to be around 300 bp apart were chosen for sequencing. Forward and reverse primers were used for sequencing as well as the T3 and T7 primers designed for the pBluescript vectors.

II.4.7.2 ExoIII deletions in the 3' end of the *ccaR* gene

In an effort to further characterize the CcaR protein exoIII deletions were targeted to the 3' end of the *ccaR* gene which would sequentially remove amino acids from the C-terminus of the CcaR protein.

The plasmid pDA150, which contains the *ccaR* gene and part of the *orf11* gene inserted into pSL1180, was digested with *Sph*I (3' overhang) and *Bgl*II (5' overhang) before being used as template for exoIII deletion reactions. ExoIII reactions proceeded as described above except that the exoIII digestion was done at 30°C rather than 37°C to slow the degradation rate and samples were removed from the exoIII mix every 25 seconds. This gave a tight clustering of deletions within the C-terminus of the *ccaR* gene. Putative deletions were screened by *Bam*HI/*Hin*dIII digestion and the most interesting clones were subjected to sequence analysis to determine the exact end of the *ccaR* gene.

II.4.8 PCR procedures

II.4.8.1 PCR of the *ccaR* coding region for expression in *E. coli*

The primer Dyl13 was designed to amplify the *ccaR* ORF and create a unique *Bam*HI site on the 5' end of the PCR product just before the start codon. The primer Dyl12 was designed to amplify the *ccaR* ORF and create a unique *Eco*RI site on the 3' end of the PCR product just downstream of the stop codon. This created a *ccaR* gene capable of creating a translational fusion with MBP or GST when inserted into pMAL-c2 (see II.3.5) or pGEX-2T (see II.3.5.5), respectively.

The reaction conditions used were: 20 pmoles of both Dyl12 and Dyl13, approximately 10 ng of the template plasmid (pDA150), 10 µl of 10X Vent buffer, 2 mM $MgSO_4$, 2% (v/v) DMSO, 2.5 mM dNTPs and 10 units of Vent DNA polymerase in a final volume of 100 µl. This was overlayed with 100 µl of mineral oil and the samples were denatured for 2 minutes and thirty seconds. The following conditions: 92°C for 45 seconds, 57°C for 60 seconds and 72°C for 90 seconds were repeated for thirty cycles using the MiniCycler PCR machine (MJ Research Inc., Watertown MA).

The extracted and precipitated PCR reactions were digested with *Eco*RI/*Bam*HI before gel purifying and inserting into pKS(+) for sequencing. The coding region of the *ccaR* clone was shown to be unaltered during the PCR and the recombinant pKS(+) clone was designated pDA180.

II.4.8.2 PCR of the *lat* coding region for expression in *E. coli*

The primer Dyl9 was designed to amplify the *lat* ORF and create a unique *Bam*HI site on the 5' end of the PCR product just before the start codon. The primer Dyl10 was designed to amplify the *lat* ORF and create a unique *Eco*RI site on the 3' end of the PCR product just downstream of the stop codon. This created a *lat* gene capable of creating a translational fusion with MBP or GST when inserted into pMAL-c2 (see II.3.5) or pGEX-2T (see II.3.5.5), respectively.

The reaction conditions used were: 20 pmoles of both Dyl9 and Dyl10, approximately 10 ng of the template plasmid (pDA172), 10 µl of 10X Taq buffer (700 mM Tris-HCl pH 8.8 and 1.0% (w/v) Triton X-100), 4.5 mM $MgCl_2$, 10% (v/v) DMSO, 2.5 mM dNTPs and 1 unit of Taq DNA polymerase in a final volume of 100 µl. This was overlayed with 100 µl of mineral oil and the samples were denatured for 2 minutes and thirty seconds. The following conditions: 92°C for 45 seconds, 58°C for 60 seconds and 72°C for 2 minutes were repeated for thirty cycles. The higher fidelity Vent polymerase would not amplify the *lat* product so Taq DNA polymerase was used instead.

The extracted and precipitated PCR reactions were digested with *Eco*RI/*Bam*HI before gel purifying and inserting into pKS(+) for sequencing. A number of the *lat* clones were shown to be mutated during the PCR procedure so the majority of the amplified DNA was replaced with wildtype DNA (from pDA172) starting from the *Nco*I site at the start codon to past the stop codon and gave the recombinant pKS(+) clone pDA182.

II.4.8.3 PCR of the pSET152 MCS for clone confirmation

The pSET152- and pDA2000-based plasmids integrate site-specifically into the *Streptomyces* chromosome using the *attP* site and integrase present on the plasmid and are therefore present at single copy. The presence and stability of the pSET152-based plasmids inserted into the *S. clavuligerus* chromosome at the ØC31 *attB* site were judged by specific amplification of the pSET152 MCS. PCR primers were generated which allowed the multiple cloning site (MCS) and insert to be amplified from the plasmid integrated into the chromosome to confirm the presence of the plasmid and the stability of the insert. The primers Dyl31 and Dyl32 were the extra-long reverse and forward sequencing primers used for cycle sequencing. These primers were shown to be identical to DNA flanking either side of the pSET152 MCS.

The reaction conditions used were: 20 pmoles of both Dyl31 and Dyl32, diluted template DNA (genomic DNA or control plasmid), 5 µl of 10X Taq buffer (700 mM Tris-HCl pH 8.8 and 1.0% (w/v) Triton X-100), 4 mM $MgCl_2$, 10% (v/v) DMSO, 2.5 mM dNTPs and 1 unit of Taq DNA polymerase in a final volume of 50 µl. This was overlayed with 50 µl of mineral oil and the samples were denatured for 2 minutes and thirty seconds. The following conditions: 92°C for 45 seconds, 64°C for 60 seconds and 72°C for 'X' minutes were repeated for thirty cycles. The length of the extension time (value of X) would depend of the size of the inserted fragment, 1 minute for every 1 kb is the rule of thumb.

In the plasmids pDA1001 to pDA1004 the 3.7 kb size of the *cob* fragment proved to be too large for amplification across the entire region so the Dyl31 primer was teamed with Dyl12 and the Dyl32 primer was teamed with Wsc13.

The amplified product was compared to the PCR products from the plasmid DNA isolated from *E. coli* on a 1.5% agarose gel.

II.4.8.4 PCR of the pDA2000 MCS for clone confirmation

The presence and stability of the pDA2000-based plasmids inserted into the *S. clavuligerus* chromosome at the *attB* site were judged by specific amplification of the pDA2000 MCS. The primers Dyl21 and Dyl33 were the extra-long primers designed to amplify the pDA2000 MCS.

The reaction conditions used were: 20 pmoles of both Dyl21 and Dyl33, diluted template DNA (genomic DNA or control plasmid), 5 µl of 10X Taq buffer (200 mM Tris-HCl pH 8.4 and 500 mM KCl), 4 mM $MgCl_2$, 2% (v/v) DMSO, 2.5 mM dNTPs and 1 unit of Taq DNA polymerase in a final volume of 50 µl. This was overlayed with 50 µl of mineral oil and the samples were denatured for 2 minutes and thirty seconds. The following conditions: 92°C for 45 seconds, 57°C for 60 seconds and 72°C for 'X' minutes were repeated for thirty cycles. The length of the extension time (value of X) would depend of the size of the inserted fragment, 1 minute for every 1 kb is the rule of thumb.

The amplified product was compared to the PCR products from the plasmid DNA isolated from *E. coli* on a 1.5% agarose gel.

III. Results. Sequencing and analysis of the cephamycin cluster

Most of the structural genes coding for the early and middle cephamycin biosynthetic enzymes from *Streptomyces clavuligerus* have been sequenced, including the *lat* (215, 234), *pcbC* (134), *cefD* (124), *cefE* (123) and *cefF* (122) genes. These genes were isolated using more classical methods that targeted only a single gene and often included a little sequence from downstream and upstream of the gene. These sequencing projects were often initiated using data obtained from the N-terminal sequencing of the target protein after purification and biochemical characterization of the enzyme. Even though the genes responsible for the biosynthesis, transport, resistance and regulation of one specific antibiotic are organized into a cluster (146), a large scale sequencing project to complete the sequence of the cephamycin cluster from *S. clavuligerus* had not been initiated. The sequence of the cephamycin cluster from *Nocardia lactamdurans* has been completed and surprisingly, the *N. lactamdurans* cluster seemed to be devoid of any open reading frame (ORF) directly responsible for its regulation.

Previous promoter probe analysis experiments demonstrated the species-specific activation of the *lat* promoter fragment in *S. clavuligerus* but not in the closely related *S. lividans*, suggesting the involvement of a transcriptional activator present only within *S. clavuligerus* (176). It was hoped that the transcriptional activator responsible for regulation of the cephamycin biosynthesis would be located in one of the many unsequenced gaps existing within the sequence of the cluster. A large scale sequencing project was initiated to sequence these gaps and complete the sequence of the cephamycin cluster.

After the sequencing project had been initiated, the DNA sequence of the transcriptional activator *dclX* (19), now referred to as *ccaR*, was provided as part of a collaboration with SmithKline Beecham and was later made publicly available in a patent (221). The gene will only be referred to as *dclX* when designating the different start codon chosen by Walters et al. (221) or when referring to the timing of its discovery with respect to the recent publication of the *ccaR*, *orf11* and *blp* sequences (174). The rest of the time it will be referred to as *ccaR*. Even though the activator had been located, the sequencing project was continued with the goal of finding all of the genes responsible for the regulation of cephamycin biosynthesis.

III.1 Cosmid mapping and subcloning of cluster fragments

The cosmid pOW309, containing much of the *S. clavuligerus* cephamycin cluster, had been obtained from P. Skatrud and provided the template for the large scale sequencing project designed to complete the sequence of gaps located within the cephamycin cluster. The cosmid pOW309 was isolated from a *S. clavuligerus* DNA genebank (124) constructed in the vector pKC462a (185) when screened with a *pcbAB*-specific probe. The pKC462a shuttle cosmid contains an origin of replication for both *Escherichia coli* and *Streptomyces* as well as bifunctional apramycin resistance (*apr*) and neomycin resistance (*neo*) markers. Insert sizes range from 26 to 40 kb of DNA inserted into a unique *Bam*HI site flanked by *Eco*RI sites within the vector.

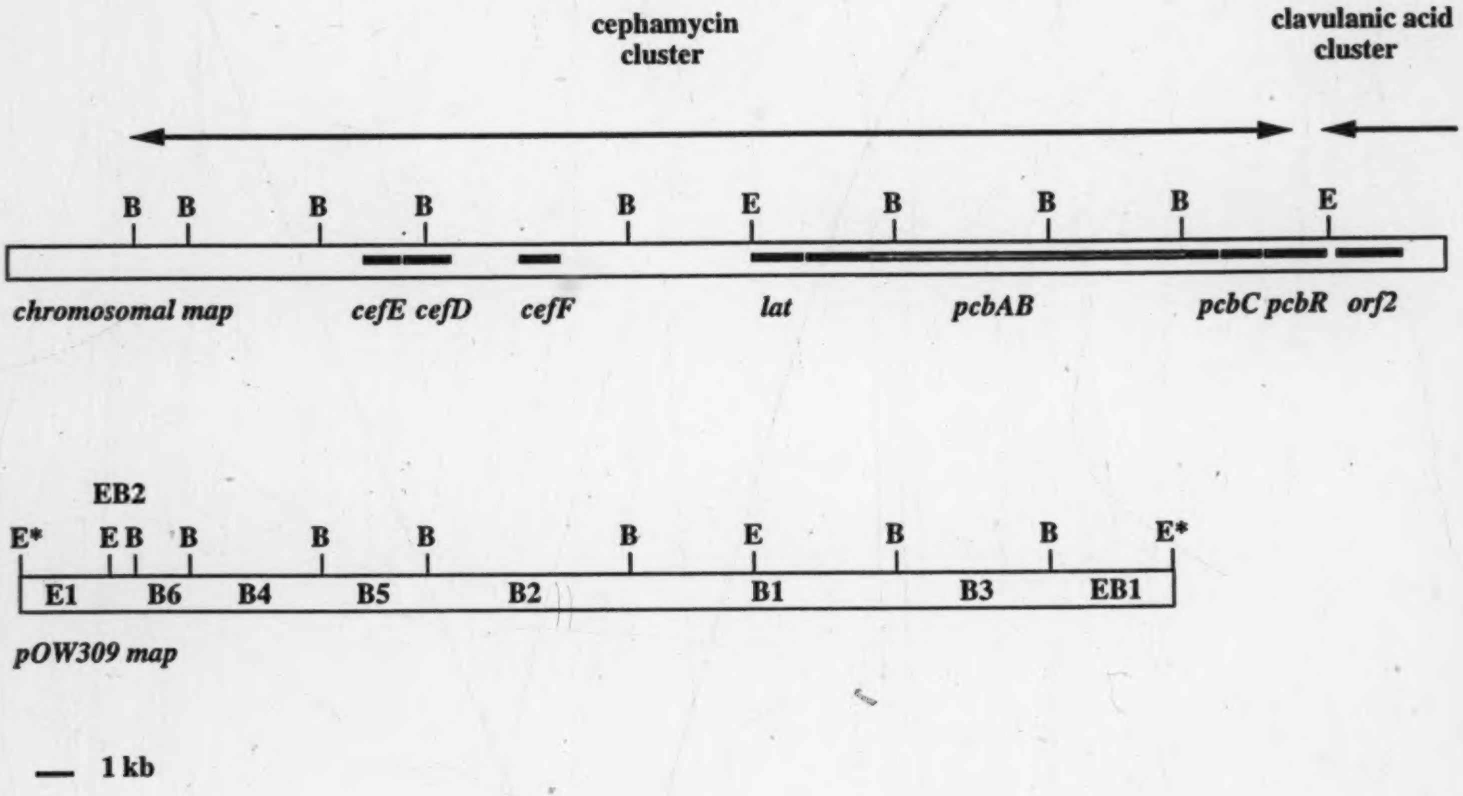
The pOW309 cosmid was mapped by Southern blot analysis of single and double restriction digests of the cosmid separated on agarose gels. The cosmid Southern blots were probed with DNA fragments containing parts of the *lat*, *pcbAB* or *pcbC* genes as well as the *Bam*HI fragments cloned from the cosmid itself. A restriction map was created (Fig. 7) and compared to a compilation of the restriction maps generated in other studies (33, 34, 142, 198, 221). The position of the already sequenced genes were included relative to the predicted location of the *Bam*HI and *Eco*RI restriction sites. The *Bam*HI fragments cloned from pOW309 were labeled B1 through B6 for *Bam*HI fragments, where B1 was the largest and B6 was the smallest. The other fragments created by *Eco*RI or *Eco*RI/*Bam*HI digestion were cloned from the cosmid pOW309 Δ*Bam*HI and labeled in the same manner. The *Eco*RI sites with the asterisks were present within the cosmid DNA flanking the insert and therefore would not be present within the *S. clavuligerus* genome.

All of the *Bam*HI fragments to be sequenced (B1, B2, B4, B5 and B6) were cloned from pOW309 into pBluescript KS(+) or SK(+) vectors well suited for the generation of exonuclease III (exoIII) deletions and sequencing. Since the *Bam*HI fragments could insert in either orientation the plasmids were digested with restriction enzymes that digest *Streptomyces* DNA frequently and are also present within the multiple cloning site (MCS) (e.g. *Sal*I, *Sma*I or *Sac*II) to determine in which orientation each *Bam*HI fragment was inserted. The plasmids were also digested with less frequently cutting restriction enzymes that recognize sites within the MCS and would give the unique 3' overhang (e.g. *Sac*I or *Kpn*I) or 5' overhang (e.g. *Xba*I or *Eco*RI) necessary for exoIII deletion reactions.

After completing the sequence analysis of the pOW309 *Bam*HI fragments an end to the cephamycin cluster had not been found. The sequence analysis from the EB2 fragment indicated that it contained a potential ORF, *sclU*, that did not show strong homology to any known sequences. The published genomic restriction maps did not include an *Eco*RI site in this location and it was initially believed that this *Eco*RI site was from the cosmid, signaling an end to the *S. clavuligerus* DNA insert and that the E1

Figure 7. Diagrammatic representation of the cephamycin biosynthetic gene cluster in *S. clavuligerus*. The chromosomal restriction map was created by compiling the restriction maps created in other studies (33, 34, 142, 198, 221). The restriction map of the cephamycin cluster containing cosmid pOW309 was deduced by Southern blot analysis and cloning experiments.

The E and B characters represent *Eco*RI and *Bam*HI restriction sites respectively and the E* are *Eco*RI sites present within cosmid encoded sequences. The solid bars represent the genes which had already been sequenced at the onset of this project. Designations B1, E1, EB1, etc. indicate the *Bam*HI, *Eco*RI and *Eco*RI/*Bam*HI cosmid fragments which were subcloned and used to prepare sequencing templates.


cephamycin cluster
clavulanic acid cluster
B
B
B
B
B
E
B
B
B
E
chromosomal map
cefE
cefD
cefF
lat
pcbAB
pcbC
pcbR
orf2
EB2
E*
E
B
B
B
B
B
E
B
B
E*
E1
B6
B4
B5
B2
B1
B3
EB1
pOW309 map
1 kb

fragment must therefore be located on the other side of the cluster. This assumption later proved to be wrong but this was only determined because of the cloning of a region of DNA which overlaps this EB2 fragment from another cosmid. The cosmid pLAFR/2-49 was originally picked out of an *S. clavuligerus* pLAFR3 cosmid (200) genebank because of its hybridization to *cefE*-specific probes (106). A *KpnI*/*Bam*HI fragment was cloned from pLAFR/2-49 into pTZ18R to provide a contiguous stretch of DNA continuing downstream from the B6 fragment. This *KpnI*/*Bam*HI fragment contained the entire EB2 fragment and most of the E1 fragment.

III.2 Sequencing of the cephamycin cluster

The double stranded DNA plasmids generated by exoIII deletion reactions were sequenced using the ABI PRISM dye terminator cycle sequencing methodologies and both strands of the fragments were completely sequenced to minimize errors. Obviously some of the *Bam*HI fragments contained genes which had already been sequenced and these genes were not sequenced again but instead were just overlapped with the new sequence. ExoIII deletion clones that would cover regions of known sequence were omitted from sequence analysis. However, in some cases a particular segment had already been completely sequenced before the published sequence became available. The B1 fragment was the largest fragment cloned, approximately 8.3 kb, but the majority of this fragment had already been sequenced as it contained all of the *lat* and *dclX* genes and some of the *pcbAB* gene. The unsequenced part of the B1 fragment could conveniently be removed as a *BclI*/*Eco*RI fragment of 2.0 kb and was sequenced before the *ccaR*, *orf11* and *blp* sequence (174) became available in the literature.

The *pcbAB* gene from *S. clavuligerus* has only been partially sequenced as an extension of the *lat* (215, 234) and *pcbC* (63) sequencing projects. Only 3.2 kb of the total gene has been sequenced or approximately 30% of the total assuming a size similar to the 11 kb *pcbAB* gene from *N. lactamdurans* (52). It was decided that the sequence of the *pcbAB* gene from *S. clavuligerus* would not be determined since little new information could be gained from this expensive undertaking over that already provided by the *N. lactamdurans pcbAB* sequence and that it was unlikely that any regulatory genes would be located in this region.

Table 5 includes the information necessary to determine how the individual fragments were sequenced and whether any special steps were necessary to complete their sequence. For example, the 6.4 kb B2 fragment was inserted into the plasmid pKS(+) and the recombinant plasmids pDA102 and pDA103 were then digested with *SacI* and *XbaI* and used as templates for exoIII deletions. The entire 4.6 kb of unknown sequence

TABLE 5. Plasmids and materials used to complete the sequencing of the cosmid fragments

Fragment	Insert size (kb)	Genes already sequenced	Total DNA sequenced (kb)	Plasmids used[a,b]	Enzymes used for 3' and 5' overhang ends	Custom primers or custom deletions
B1	8.3	*dclX*, *lat* and *pcbAB*	2.0	pDA122	*SacI*/*Xba*I	no
	8.3	*dclX*, *lat* and *pcbAB*	2.0	pDA123	*SacI*/*Xba*I	no
B2	6.4	*cefD* and *cefF*	4.6	pDA102	*SacI*/*Xba*I	Dyl17/Dyl23
	6.4	*cefD* and *cefF*	4.6	pDA103	*SacI*/*Xba*I	Dyl14
B4	4.1	no	4.1	pDA107	*SacI*/*Xba*I	no
	4.1	no	4.1	pDA107	*KpnI*/*Eco*RI	pDA124
B5	3.2	*cefD* and *cefE*	1.6	pDA109	*SacI*/*Xba*I	no
	3.2	*cefD* and *cefE*	1.6	pDA110	*SacI*/*Xba*I	no
B6	1.7	*bla*	0.6	pDA111	no	pDA125
	1.7	*bla*	0.6	pDA111	no	pDA126
EB1	0.8	no	0.8	pDA115	no	no
	0.8	no	0.8	pDA116	no	no
KB1	2.8	no	2.8	pDA119	*SacI*/*Acc*65I	no
	2.8	no	2.8	pDA119	*SphI*/*Bam*HI	no

[a] Parental plasmids used as the template for sequencing, carrying out exoIII deletion reactions or preparing custom deletion subclones.

[b] The sequentially numbered plasmids represent opposite orientations of the same inserted fragment.

was determined on both strands using the exoIII deletion clones except for regions in which deletion subclones could not be found so custom sequencing primers were designed. The custom primers were designed from new sequence data acquired to initiate sequencing just before the unsequenced region. Custom deletion plasmid which resulted from use of unique restriction sites within the new sequence and the MCS that would put the region of unknown sequence next to the primer binding sites, were also used.

The junctions of the fragments were sequenced using clones which overlap the *Bam*HI site or other junctions. This was done to prevent the exclusion of any sequence if, for instance, two *Bam*HI sites were located very close together. The junction of the *Bcl*I site between the sequence provided by Walters et al. (221) and that of the *BclI*/*Eco*RI fragment in pDA122/pDA123 was sequenced using the pDA130 plasmid initiating at a *Nco*I site near the *Bcl*I junction. The *Eco*RI junction between the sequence of the *BclI*/*Eco*RI fragment in pDA122/pDA123 and the published sequences of Tobin et al. (215) was sequenced using the pDA128 plasmid initiating at a *Nco*I site near the *Eco*RI junction. The *Bam*HI junction between the B1 and B2 fragments was already determined by Walters et al. (221). The *Bam*HI junction between the B2 and B5 fragments was already determined by Kovacevic et al. (123). The *Bam*HI junction between the B5 and B4 fragments was determined using the custom primer Dyl22 with the pOW309 Δ*Kpn*I cosmid template. The *Bam*HI junction between the B4 and B6 fragments was determined using the custom primer Dyl24 with the pOW309 Δ*Kpn*I cosmid template. The *Bam*HI junction between the B6 and EB2 fragments was determined using the custom primer Dyl28 with the pOW309 Δ*Kpn*I cosmid template.

III.3 Organization of sequence data from the cephamycin cluster

The sequence data in Fig. 8 were compiled from the sequence information determined in this study as well as that previously available from other studies (see below). The figure includes the sites for the restriction enzymes used in cloning or generating mutations throughout this work. The restriction enzymes *Eco*ICRI and *Acc*65I are isoschizomers of the enzymes *Sac*I and *Kpn*I, respectively and since they recognize the same restriction sites only the *Sac*I and *Kpn*I sites are indicated. The ORFs are designated by uppercase text and their orientation and gene designation are also shown. The organization of the ORFs was determined using a combination of the CODONPREFERENCE program (60) from the UWGCG package, which identifies ORFs on the basis of the characteristic third-position GC bias typical of *Streptomyces* (25), as well as DNA Strider 1.2 (144) and BLASTX and BLASTP database searches (9).

Figure 8. Compiled sequence data for the cephamycin cluster of *S. clavuligerus*. The sequence results generated from this study and those from other studies were assembled into a single contiguous sequence spanning from *rhsA* to *pcbAB*. Sequence containing ORFs were designated by upper case text while the intergenic regions were given in lower case text. Cosmid subclones fragment designations were underlined, while the gene designations and the arrows indicating the direction of transcription were printed in bold. The sites for commonly used restriction enzymes and the total nucleotides of sequence were also listed.

The nucleotide sequence for the *bla* ORF (nts. 3147 to 4257) was described in Perez-Llarena et al. (173) with the nucleotide accession number of Z54190. The nucleotide sequences for the *cefE* and *cefD* genes (nts. 10181 to 12675) were described in Kovacevic et al. (123, 124) with the nucleotide accession numbers of M24140 and M32324. The nucleotide sequence for the *cefF* gene (nts. 14951 to 15908) was described in Kovacevic and Miller (122) with the nucleotide accession number of M63809. The nucleotide sequence for the *dclX* gene and upstream sequence (nts. 18172 to 19917) was provided by B. Barton (19) and was described in Walters et al. (221) but does not have a nucleotide accession number. Differences were noted between the two *lat* sequences (215, 234) and only 141 bp of sequence (nts. 21867 to 22008) was from Tobin et al. (215) with the nucleotide accession number of U12015 was included in the compilation. The remaining nucleotide sequence for the *lat* gene and the *pcbAB* gene (nts. 22009-26035) was described in Yu et al. (234) with the nucleotide accession number of M64834.

```
KpnI (1)
GGTACCCGGCTACCCCGAGAGCGGTGCCCGCACCACCCGCACCGCCTACGACTGGTCCACCGGGCTGCCGCTCTCCAGGA  80
 KB1     •rbsA--> •         •         •         •         •         •         •
                                                SacII (129)       NruI (147)
CGGAAGCGCCCGACAGCCTCGCACTGAGCCGTACCACCACGTACAACGCCGCGGGACAGCCGGTTCTCGCGACCACCCCG  160
         •         •         •         •         •         •         •         •

GCCTCGGACGGCCAGGACGCCTCCGCGACCGTCACCGCCTACTGGACGGCAGGCGGCACGGGCGACTGCTCCGGCCGCCC  240
         •         •         •         •         •         •         •         •
                                SacII (273)
CGAGTGGGCCGATCTGGTCTGCCGGACCAGCCCCGCGGCGAAGATCACCGGAGCCAACGGCAACCCCGACGAGGGCCTGA  320
         •         •         •         •         •         •         •         •

CCACCACCACCGAGTACAACCGGCACGGGCAGACCACCAGGACGACCGAGACCGGCAACGGCACCACCCGCGTCACCGAG  400
         •         •         •         •         •         •         •         •

ACCGGCTACGACGACGCCGGCCGGGTCACGCGGACAGCGATCACCGGCGACGCCGGGGCGGCCGTCGAGGAGAACACGAC  480
         •         •         •         •         •         •         •         •

CGGCTACGACCGGGAGAACGGCAGGGTTGTCGCGGTCTCGACCGCCGACGCCACGTTGACCAGCGCCTATGACGCCCTGG  560
         •         •         •         •         •         •         •         •
                             NarI (590)
GGCGTCTGGTCGGATACACCGACGCGAACGGCGCCACCACCGCCACCGCCTACGACGCCCTGAACCGGGTGGTCACGACC  640
         •         •         •         •         •         •         •         •
                                                                   StyI(708)
ACCAGCTCCGCACCGTACACCCTCACCTACACCTACGACACCGCCGCCGAGCCCCGTGGCCTGGCGACCAAGGTGACCGA  720
         •         •         •         •         •         •         •         •

CTCCGTCGCCGGTGAGTTCACCAGCGTCTACGACGCCGACGGACAGACCGTCCGGCAGACCCTTCCCGGCGGCGTCACCA  800
         •         •         •         •         •         •         •         •

TGCGCCAGACCAACGCGCCCAACGGCGAGGCGGTCGCCCGGAGCTACACCCTCGACGGCCAGGAGGAGCCCTTCTTCTCC  880
         •         •         •         •         •         •         •         •

AGCGCCGCCACCACCACGGTGCACGGTGAACTGGCCGAGCGGACCGGTCTGACCGTGGCCGACAACGCCTACGACGCGGC  960
         •         •         •         •         •         •         •         •

GGGCCGTCTCGTGCGGAGCGCCACCACCGCCCAGGACGGCTCCGCCGACTGCGCCCTGTACAGCTACACGTACAACGCCA  1040
         •         •         •         •         •         •         •         •

ACGGAAGCCGCAAGAGCCAGGCCGCCGGGACCGGCAACGCCGAGACGGGCTGCCCCTCCGCCACCGGCACCCCGACCGAG  1120
         •         •         •         •         •         •         •         •
                           SalI (1148)
CACGTCTACGACTCGGCGGGGCGCATCGTCGACGACGGCTACCGCTACGACGGACGGGGCCGTCTGGTCTCCTCCCCGGA  1200
         •         •         •         •         •         •         •         •
              ScaI (1215)             NarI (1239)
CGGGCTGTCCTCGGAGTACTTCGCCAACGACACGACCCGGCGCCAGACGGCCGACGACCGCCGGCAGACCTGGACGCTCG  1280
         •         •         •         •         •         •         •         •
                                          SacII (1323)ApaI (1335)
ACCCGGCGGGCCGCAAGCGCGGCGTCCTGATCGAGACGAACACCGCGGGCGCATGGGCCCAGACCGATGCCGAGCAGTTC  1360
         •         •         •         •         •         •         •         •
                                                                    (1436) SalI
CACTACGACGAGAACGACCAGCTCGTCTGGGTGCGGAACGGCAGCGGGGGCGTCACCCGCACGCTGGTGACCGACGTCGA  1440
         •         •         •         •         •         •         •         •
                    PvuI (1461)             PvuI (1485)
CGATCTGAACGGCATCGCGGCGATCGGCGACGACGGCTCGGTGACGATCGAACTCAACGAGGCCAGTGACGAGGGACAGA  1520
         •         •         •         •         •         •         •         •

TCCAGTACACCCCCGGCACCGGGGACGTCCTCGTCCAGGACAGCGCGGCCGGACAGCCCGCCGGAGCGGATCTGGCGGCT  1600
         •         •         •         •         •         •         •         •

CTGATCGGCGGCGAGGACGTCACCGGCGCGGCCGGTGCCCTCGCCCGGTGCCTGGGCTGCGTGGCCTCCCCCCTCGCCTC  1680
         •         •         •         •         •         •         •         •
           SacII (1692)                                                 SacII (1753)
GCCGATGAGGGCCGCGGCGCGCTACTCGCCGAAGGGTTCGGCGAAGTCCGCCACGGACTACTGGCTCTTCACCCGCGGCC  1760
         •         •         •         •         •         •         •     KB1 •
```

```
                                      SacII (1798)
TGTGGGCCTTCCGGGACGAGAAGCGCCGCAAGGGACGCCGCGGTGACCTCGTCTGGAGTGACGACGGGTGCTCCGCCCCG  1840
 KB1     .         .         .         .         .         .         .         .

TGGTACTCGCACATCGTGATCGGGCCTTCGGTCGGCTACTACAGCGGTCAGTTCTACTGGCCCTGCGCCCGGCACGACTT  1920
         .         .         .         .         .         .         .         .

CGGCTACCGCAACTACCGGAAGCAGAACCGTCGCACCAGGGCGAACAAGGACAAGATCGACAACCGGTTCCGCTACGACA  2000
         .         .         .         .         .         .         .         .
                     StyI (2022)EcoRI (2033)
TGAAGAAGCGGATCTGTGAGCCCAAGGCCGTCGAATTCGTGTGCAAGGGGGCGGCGGACGGCTTCTACTACGCGGTGAAA  2080
         .         .         .         .         .         .         .         .
                                              PvuII (2136)
AAGGGTGGGGACAGCTCCTTCTTCAAGAAGTGGTGAccctcgccgacatgagaaccagctgatcgctccgtggccggtca  2160
         .         .        rhsA-->   .         .         .         .         .

ggggggtggaccgagaggaacactcccctgaccggccacacccgaccccgcctgaggatcgccccccATGAACAGAACGT  2240
         .         .         .         .         .         .        sclU-->    .
                                 SacII (2274)
CACACCCCTCGTTCCGTATCGCCCTCTCCCTCGCCGCGGCCGTGTGCTCGGCCGTCGCCGGACTCGTACTGCTGCCCTGG  2320
         .         .         .         .         .         .         .         .

GACCCGAGAGCCACGGTGCCGGTCGCCCCCGGATCGGACACCCTGACCACCGGGTACACCTGGTGGGGGGTCGCGCTCTT  2400
         .         .         .         .         .         .         .         .
                                                       NarI (2456)
CGCGGTGGTCCTGGCGGTCGCCGCCTGGGTCACCGGCCGTTTCGGCGGTGGCGCGGGCGCCGCCATGGCGGCGGCGGGCG  2480
         .         .         .         .         .         .  NcoI/StyI (2463) .

GCGTTCCCTCCGCCCTGCTGCTGGTCTCCTTCCTGGCGCATGAACCCGAATCCGACGGCTTCGCGGGGCTCTGGTCCCTC  2560
         .         .         .         .         .         .         .         .
                                 NarI (2594)  NruI (2607)
AACTGGCTCTTCGCCGCCTGTCTCATGGTCGTCGGCGCCGGAGTCGTCGCGATGGTCGCCCAGGACGGCCGCCCCGGACT  2640
         .         .         .         .         .         .         .         .
                                         SacII (2682)
CGACCGGATGCCGGATGCCTCCCTGAAGCTCCTGGTCCCGGCCGCGGTCGCGGGAGCGGTGTCGTTCACCGCGTTCCTCG  2720
         .         .         .         .         .         .         .         .

TCCAGCTCGTCTCCTCCGGTTCGGGGCCGGTGGTACGCGGCTGCGCCGCCCTGGCCGCCCTCGCCCTCGGGTACGGCGCG  2800
         .         .         .         .         .         .         .         .
    BamHI (2805)
GGCTGGATCCTGCGGGCCGTCGGCCGCCGCCCGGCCACCGCGACCCTGCTCACGGTCGGTGCCGTGCTGCTCATGGCGGC  2880
<--KB1  B6-->      .         .         .         .         .         .         .
                                    NarI (2917)
CGTCGAGACCGTATCCGCGCTGCTGTGGGAGTGGCGGGCGCCGGAAGCCCTCGCCTGGGGGATCACCCTGCTCCGCTGTG  2960
         .         .         .         .         .         .         .         .

GAGCGGGCACGGCGGCGGCGTATCTGCCGCGCACCCCGAACGGTCCCCCGTCGGGGGCGTGAcggacggccggagacccg  3040
         .         .         .         .         .     sclU-->     .         .
                                        NruI (3081)                SmaI (3108)
caagcggcctcctccggcttccgctccgtgcgccgcgccgtcgcgacaccacggcgttaccgctgtacccgggcgcgtac  3120
         .         .         .         .         .         .         .         .
                          BclI (3147)
cggtagcgtgcgtggccccacgttcgtgatcacagagcgcccggcgcggctccgcgccgggtggtgagagcaggtgacgg  3200
         .         .         .         .         .         .         .         .

cgcatggtcttcgagcggtcccggcggccggagggacctgtccgggcggagggacccGTGCGGGCGGGGGCTGTCCGCCC  3280
         .         .         .         .         .       bla-->      .         .

GTCCGCCGTCTCCCGCCGGGCGCTGCTGACGGCAGCGGCGCTGCTGCCGCTCGCCGGGTGCGGTACGAGGGAACGGACCG  3360
         .         .         .         .         .         .         .         .
                                            NarI (3405)
GCTCCCCGCCCGCGTCGGTACGGGCTGCGGGGCGCGGGTCGCCCGGCGCCGAGGGGCGGGCGCGGGTGGCGGACCTGGCC  3440
         .         .         .         .         .         .         .         .
                    ApaI (3460)
GCGCTGGAGCGGGAGCACGGGGCCCGGCTGGGGGTGTACGCCCTGGAGACCGGCACCGGGGCCGAGGTCGCCCACCGCGC  3520
         .         .         .         .         .         .         .     B6 .
```

```
     Eco47III (3526)                          SacII (3568)
CGACGAGCGCTTCGCCTTCTGCTCCACGTTCAAGGCCCTCGCCGCCGCCGCGGTCCTGCACCACCATCCGATCCGCCATC  3600
 B6      .         .         .         .         .         .         .

                      SmaI (3622)SalI (3633)
TGGAGCGGCGGGTCACCTGGACCCGGGCCGACGTCGACTCGATCTCACCGGTCACCGAGGACCACATCGCCACCGGGCTG  3680
         .         .         .         .         .         .         .         .

                                                                   (3759) ApaI
ACGGTCGGCCAGCTCTGTGACGCCGCCATCCGGCACAGCGACGGCACGGCCGGGAATCTGCTGATGCGCGATCTCGGCGG  3760
         .         .         .         .         .         .         .         .

GCCCTCCCGGCTCACCGCCTATCTGCGCGGACTCGGGGACTCCGTCAGCCGGATGGACCAGTACGAGCCCGAGCTGAACC  3840
         .         .         .         .         .         .         .         .

                  SmaI (3859)     SmaI (3877)
ACGACCCGCCGCACGACCCCCGGGACACCACCACCCCCCGGGCGATCGCGTCCGACTACCGGAAGCTGGTCCTCGGGGAC  3920
         .         .         .         .  PvuI (3883)        .         .         .

GCGCTGACCCCGGACCGGCGCGCCCTGCTCACCGACTGGCTGGTGCGGAACACCACCGGAGGCCGCCGTATCCGGGCCGG  4000
         .         .         .         .         .         .         .         .

AGTCCCCTCCGGCTGGCGGGTGGCCGACAAGACCGGAACCGGCAACTACGGCAGGGCCAACGACATCGCCGTGCTCTGGC  4080
         .         .         .         .         .         .         .         .

CGCCCCGGTCGTCCCCGCTGGTAGTAGCCGTCATGTCCGACCGTCCCGGCTTTCGCACCCCGCCCAGCGAACGGCTGATC  4160
         .         .         .         .         .         .         .         .

        SacII (4169)                                                NarI (4229)
GCCGAGGACCGCGGAGCGATCGTCGCCGGGCTGGTCCGGCACCCGCCGACCCATGCGTCAGGGGCCGGGGCGCCCTCGCG  4240
         .      PvuI (4177)    .         .         .         .         .         .

                                        ApaI (4278)
GCACCCCGGCCCCTGAcgggacggaacaggacgagccgggcccgaggggcggTCAGCCGAGCCGGGTCAGCCTGGTGCCC  4320
         bla-->    .         .         .         .  <--pbpA          .         .

AGGGACGGCTCCGTCAGCCCCTCCTTGAGGGCCAGCGGAATCCGGTGCGCGTCGGGACCGTTCCCCACCGTCAGCTCACC  4400
         .         .         .         .         .         .         .         .

GACCACGGCCCCGGCCCTCGCGGTGCGCGAGAGGGCTCCGTCGCCCTTGCGGCCGTAGGTCAGCGTCACCTTCTGACCGG  4480
         .         .         .         .         .         .         .         .

                                        BamHI (4520)          SalI (4542)
ACAGGCCGACCAGATCGAGGGACTTGGTCGTCACCACCGGGATCCGGCGGCCGAAGCGGTCGTCGACATGGCCGACGACC  4560
         .         .         .    <--B6     B4-->            .         .         .

TCGCCCTTCTTCACCACGTTCGCGGAGGTCAGCATCTCGCGCACCGAGGCGACGACCTTCCTGCTGTTGTCCTGCACCAG  4640
         .         .         .         .         .         .         .         .

                                  ApaI (4668)
CACCAGGCTGTCGCCGCCGTTGATGTCGGGCCCCGGGGCGTGCTGGGCCATCATCGTGCCCAGGATCAGCCGGTCCTCGC  4720
         .         .         . SmaI (4672)        .         .         .         .
                                        SacII (4760)
CGTCGAACGTCTTGTACGCGGCCCAGGACAGCGTGCCGCCCGCGGGGGTGTTGGAGCCGGTCTTGATGCCCACGATCCCC  4800
         .         .         .         .         .         .         .         .

AGACCGGCCAGTATCAGCCGGTCGTTGTTGTTGTAGATCCGGCCGTGGCCGGGGATCTCGGCGTTCGGCAGGGCGACGAC  4880
         .         .         .         .         .         .         .         .

                                                SalI (4929)
ATCGCGGAAGACGTCGTACTCCATGACCGCCTCGGCCAGCTTGAGCTGGTCGACGGCGGTGCTGACCGTCCCCTTGTCCA  4960
         .         .         .         .         .         .         .         .

GGCCGCTGGGGTCGGTGTAGACGGTGTTCTTCATGCCGAGGTCCTTCGCGGCCTCGTTCATCTCCCGGACGAACGCCGTC  5040
         .         .         .         .         .         .         .         .

TGATCGTCCGTCTTGGTGGACCAGCGGGCCAGCAGCCGCCCTATGTTGTTGCCGGACGGGATCATCAGCATCTTCAGCAT  5120
         .         .         .         .         .         .         .         .

                                                       StyI (5174)
GTCCTGCTGGCTGAAGGTCTGACCCTCCGTGAGTCCCTCGATCCGGGACTCGTCCTTGGCCTTGCCCTCCGCCACGGTCT  5200
         .         .         .         .         .         .         .         .

    SalI (5205)       SmaI (5223)
GCGCGTCGACCGTGATGGTCGGCCCGGGTTCGCCCCGCTTGAGGGGGTGCTTCCGCAGCAGCACATAGGCCGTCATGATC  5280
         .         .         .         .         .         .         .      B4 .
```

```
TTGGCGGTGCTCGCCGTCGGTACGGACTTCTCCGGGCCGAAGGTCTCGTAGACGCCGGAGCCCGCGACGGTGACCGCGCC  5360
B4       .         .         .         .         .         .         .         .
        StyI (5369)
CTGTCCCTCCTTGGGCCAGGGTATGGAGATCCGGTCCCCGCCGAGCGTGGTGGACGAGGCGCTGTCGCGCAGGGTGAGGA  5440
         .         .         .         .         .         .         .         .
                                                                        NruI (5513)
CCGGGTCCGGGAGCGGGCGCAGGAGTTGGACGGCGACGAAGGCGATGACGAGCAGCAGCGCCGCCAGCGGGATCGCGATC  5520
         .         .         .         .         .         .         .         .

ATggccaggcgggcgcggccacgggagcgggggcggcggacgacccggcggggacggagagccgggcgagcagatcgag  5600
<--pbpA  .         .         .         .         .         .         .         .

ccgcgcgcgggcgtcctcccgttcctgctcccgcttccgggtctgctcctgccgctgccgctggaccagtgccgcgtccg  5680
         .         .         .         .         .         .         .         .
           NarI (5692)
aggccgcgtccggcgccgacgcggcggcgggtgtccgggcaggcttcggctccggtacgggcttcggcttcggggccttg  5760
         .         .         .         .         .         .         .         .

tcccgctccggtgtcctgtcccgctccggggcctgcttctgtgcctgcttctgtgcctgcttctgtgcctgctccggtgt  5840
         .         .         .         .         .         .         .         .
      SacII (5847)
ctgctcccgcggcgtgggcacgggcttcgcgggcgcgtccgggtccacggcccaggacgggcgcggcttgcgcagactgc  5920
         .         .         .         .         .         .         .         .

gggtggcgtcgtccggaacgggagcgtccgaggtcgggtccgcgtccttcccgtccgtgcccgtgcggtccgtgcccgtg  6000
         .         .         .         .         .         .         .         .

cgcggcagcgcgatcacccgcgtggtctccgaggactcggcgggcttcgcgggctcggccgtggcctccggctccggctc  6080
         .         .         .         .         .         .         .         .
           StyI (6092) SmaI (6099)
cggcttcggctccttgggcccgggaacggcggactccggcttctcgtcagccggagtctccgcctcctcagccttttccg  6160
         .     ApaI (6096)   .         .         .         .         .         .

acccctctgtgccctcggcggactcggcggactcggcgtccaccgtctcgtcgtcggccccgtctcgtcgtccgcgtcc  6240
         .         .         .         .         .         .         .         .

tcgctccggggcggcacctcgtcggtggccggtgaggtcgtcccctcgtcggccgggcgcggatcgcgttcgtcgccgtc  6320
         .         .         .         .         .         .         .         .

gttccccgtcggcacgttcttctggtcgagtgtgtcgggggaccgccccgccattgctgtctccttcatacgccgcgccg  6400
         .         .         .         .         .         .         .         .
                                                                    (6477)  SmaI
tcgctctctgccccacgtggccgtgtcgtctgtccgggcgtgccccgtgccgaaacccggccccggacgcgctccccccg  6480
         .         .         .         .         .         .         .         .
                                          PvuI (6523)               PvuI (6549)
gggtgcggacaccgttcgctctgcggcggcgaccagcgtaaccgatcggtgttcgagggtgagcgggccgatcggaccgt  6560
         .         .         .         .         .         .         .         .

accggcggtgcctccttccgctggtacgggggagttgcggcaccccgttcccgagcggccggagcctgtcggcccgctc  6640
         .         .         .         .         .         .         .         .
      SmaI (6647)
tcccgccccgggggaagacggcaccgggccgcacctggttgcagacgccggacagcgggcggtgggcggacggcggatgg  6720
         .         .         .         .         .         .         .         .
                                                                NarI (6785)
cggacggTCAGCGAACGGCGGGGCGGGTGTCGCCGTCGGCCCCGTCGGTGTTCTCCCCGGCCCCGGCGCCGGGGCGGGCG  6800
         .         .         .         .         .         .      (6797)  NarI

CCCGACCGGCTGTCGCGCAGCGCGGGGAGCAGGGGCGTGAGCGCGGCGAGGACGAGCAGGATCACGGCCATCGAGAGGAA  6880
         .         .         .         .         .         .         .         .

GGCGTCCCCGTAACTCTCCGCCAGCGCGGCGGGGCTCCCGGCCGCGCCCGTGGTCCCGGTCGCCGTGAGGGTGAGAACGA  6960
         .         .         .         .         .         .         .         .

CGGCGAGGCCGAACGCGCCGCCGAACTGCCGGGTGGTGTTCATCAGGCCGGAGGCGGCCCCGGCGTCGGCGTGGCTGACG  7040
         .         .         .         .         .         .         .   B4  .
```

```
CCGGAGGTGACCGTCCTGGTCAACGGGGTGCTGAGCAGACCGCCGCCGAGCGAGAGCAGCACGGCCGGGCCGAGGATGCC  7120
 B4      .         .         .         .         .         .         .         .
                                                           PvuI (7180)
GGTGACGTAGCCGCTGTCGGGGGTGATGGCGCTCTGCCACCAGAAGCCCCCCGCCTCCACGATCGCCCCGGTGGCGATGA  7200
         .         .         .         .         .         .         .         .

GGACCCTGGCCTGGACCCGGCGCATCAGCAACGGTATGAGCCGGAGCGAGACCAGCAGCATGACCGCCGCGTGCGGCAGG  7280
         .         .         .         .         .         .         .         .

AAGCCGAGTCCGGTCTGCGCGGCGCTGTAGCCCAGGACGTCCTGCATGGTCAGGGTGAGGAAGTACCAGAGCGGCACCTG  7360
         .         .         .         .         .         .         .         .
                                          SmaI (7403)
GCTCGCCCCCGCCAGCAGCATCGCCAGATTGCCCCAGCCGATCCCGGGCAGCCGGAACAGCCGGGGCGGGATCAGGGGCG  7440
         .         .         .         .         .         .         .         .

TGGCGGCGAAACGCGCCTCGACCACCGCGAACGCGACGAGGGCCGCGCCGCCGCCGAGCAGCGGCAGGGCCGTGCGGGGG  7520
         .         .         .         .         .         .         .         .

TCGGACCAGGCGTGGGTGTGGGCCTGGGCGACCCCGAAGGTGAGCAGGGTCAGCGCGGCCGTGGCCAGGACCGCGCCCGG  7600
         .         .         .         .         .         .         .         .
                         NarI (7626)                             SacII (7666)
CAGGTCGAGCCGGACCGCGCGCTCCGGCGCCCCCCGGCGGGCGAGGACCCGCAGGCCGAAGAGGGCCGCGGGAATCCCGA  7680
         .         .         .         .         .         .         .         .

CCGGCACATTGACCAGCAGCACCGAGCGCCAGGAGAGATATTCGGTCAGCAGCCCGCCGAGCAGATTGCCCGCCGCTCCG  7760
         .         .         .         .         .         .         .         .
             SacII (7774)     ApaI (7792)      ApaI (7809)
CCGGCCAGGGCCACCGCGGTCCAGATGGTCAGGGCCCTGGTACGGCTGGGGCCCTCGGGGAAGCTGGTCGTCAGCATGGT  7840
         .         .         .         .         .         .         .         .
                        NarI (7865)         SmaI (7885)                   NruI (7915)
CAGCGACAGGGGGGCGAGGACGGCGGCGCCGAGGCCCTGTGCGGCCCGGGCCGCGATGAGGGTCGCCGGGTCGGTCGCGA  7920
         .         .         .         .         .         .         .         .

GCCCGCCGAGCAGGCTGGCGACGGTGAAGACGGCGAGGCCCGCCGTGAAGACCCTGCGGTGTCCGTACAGGTCGGCCAGG  8000
         .         .         .         .         .         .         .         .

CGTCCGCCGAGCAGCAGCAGTCCGGCGAAGGCCAGGGTGTAGGCGCTGACGACCCATTGGAGGCCGGAGGGGGTGAACCC  8080
         .         .         .         .         .         .         .         .

CAGGTCGGACTGGATCGACGGCAGCGCCACGGTCATCACCGAGACGTCGAACAGGACCAGGAAGTGCGTCGTACACGACA  8160
         .         .         .         .         .         .         .         .

GCGCGAGCGCCGTCCTCGCAGCGGAACTCCCTCGCATGGCTGGGCCTCTCCCCCGTGTGGCGTGTATGAGGTCCGCGCTC  8240
         .         .         .         .         .         .         .         .
                                                          NarI (8299)
CGGCGGGCGCGGGGTGCTGCGGACACgcgcgggtgcgccccgcctgccctgggtgcggggcgccacggcgtcatccgtgc  8320
         .        <--cmcT    .         .         .         .         .         .
SphI (8322)
ggcatgcccccccgtaggcccccctgaggacccccggaggaccgggccgggacgagtcaggacgtcccggacggccgggtc  8400
         .         .         .         .         .         .         .         .

cgggacgcgccggtgccggacgagccggtgcgatgtgccccccgtacgacgcgtgcgggccgtgccgcccgcccccgccc  8480
         .         .         .         .         .         .         .         .
                                                                   (8556)  SacII
ctgtccccgaccggcccggaacggcagggcagttcgctcccccggcgcggcggcgggcgcggcaccgggggtcgtccgcg  8560
         .         .         .         .         .         .         .         .
   NarI (8563)     BamHI (8579)       NarI (8599)SacII (8610)
ggggcgccgtcccgtgcgggatccctcccgcacgggacggcgcccggcgccgcgggccggggcggaCTACTGGGAGAAGT  8640
         .    <--B4     B5-->.         .      NarI (8606)  .       <--pcd      .

CCACGCCCTGGGCCAGTGCCACCCGCCCGGAGTAGTTCACCGTGTTGGTCGCCCGGCGCATGTACGCCCGCCAGGCGTCG  8720
         .         .         .         .         .         .         .         .
            NotI (8732)
GAGCCGGACTCGCGGCCGCCGCCGGTCTCCTTCTCGCCGCCGAACGCGCCGCCGATCTCCGCGCCCGAGGTGCCGATGTT  8800
         .         .         .         .         .         .         .      B5 .
```

```
                                         Eco47III (8842)
GACGTTGGCGATGCCGCAGTCCGCGCCGTCGGGGGCGAGGAAGCGCTCGGCCTCGCTCTGGTCGGCGGTGAAGATCCCTG  8880
 B5      •         •         •         •         •         •         •         •

CCGACAGGCCCTGCGGGACCTCGTTGTTCAGCCGGATGGCCTCGTCCAGGTCCCGGTAGGTGAGGACGTACAGGATGGGC  8960
         •         •         •         •         •         •         •         •
                                                                  NarI (9029)
GCGAAGGTCTCCTCGCGGACCACGGCGGTCTGCGCGGGCATCCGGACCAGCGCGGGCCGTACGTAGTAGGCGCCGGGCGC  9040
         •         •         •         •         •         •         •   NarI (9036)

CGCGTCGGGGAACTGCCGCTCACCGCCCGCGCACAGGGTGCCGCCCTCGGCGGTCGCGCGCTCGACCGCCTCCCGCATCC  9120
         •         •         •         •         •         •         •         •
                                                            PvuI (9181)
GGCCGAAGGCGGCCTCGTTCACCAGCGGTCCGACCAGGGTGGTGTCCTGGAACGGGTCGCCGATCGGCAGCCGCTCGAAG  9200
         •         •         •         •         •         •         •         •
                                                                      Eco47III (9271)
GCGGCGGTGAGCCGCTCGACGACGGTGTCGGCGATGTCCTCGTGCACGATCAGCCGACGGAGCGTGGTGCAGCGCTGGCC  9280
         •         •         •         •         •         •         •         •
        Sac II (9289)
CGCCGTGCCCGCGGCGGCGAAGACGGCCGCGTTGACGGTGAGGTCGAGATCGGCGGACGGCGTCACGACGGCGGCGTTGT  9360
         •         •         •         •         •         •         •         •
                                          SmaI (9403)
TGCCGCCCAGCTCCAGGATGGTGCGGCCGAACCTGGCGGCCACCCGGGGGCCCACGGCGCGGCCCATACGGGTCGATCCG  9440
         •         •         •         •       ApaI (9408) •         •         •
                  SmaI (9460)
GTGGCGCTGACCAGCGGCACCCGGGGGGAGTCAACCAGCCGCTCGCCCACGTCGGCCGCGCCGACGACGACCTGGTTGAG  9520
         •         •         •         •         •         •         •         •
       NarI (9528)
GCCCTTCGGCGCCCCGGCGTCGGCGATGGCGAGGTCCAGCAGGGCGGCGCACGCCAGCGCGGTCAGCGGGGTCAGCTCGG  9600
         •         •         •         •         •         •         •         •

ACGGCTTCCAGACCACGGTGTCACCGCAGACCAGGGCCACGGCCGCGTTCCACGCCCAGACGGCGACCGGGAAGTTGAAG  9680
         •         •         •         •         •         •         •         •
    BclI (9685)
GCGCTGATCACACCGACCACGCCCAGCGGGTGCCAGGTCTCCATCAGCCGGTGGCCGGGGCGCTCGGAGGGCATCGTCCG  9760
         •         •         •         •         •         •         •         •

GCCGTAGAGCTGCCGGGACAGGCCCACCGCGAAGTCGCAGATGTCGATCATCTCCTGGACTTCGCCGAGCGCCTCGGAGC  9840
         •         •         •         •         •         •         •         •

GGATCTTCCCCGCCTCGATGGTGACCAGGTCGGCGAGGTCCTGTTTGTGCTCGGTGAGCAGTTCCCCGAACCGCTTGACC  9920
         •         •         •         •         •         •         •         •
  NarI (9923)
AGGGCGCCGCGCACCGGGGCCGGGGTGGTGCGCCAGGTGAGGAACGCGGTGTGCGCGGCCTCGACGGCGCGGTCCACGTC 10000
         •         •         •         •         •         •         •         •
                                                                    (10078) SacII
CTCGGGGGTCTGGGCGCGCAGGCCGAACAGATCCGCGCCGGTGAGCGGGGTGCGGGCCGTGAGGGCGTCGCCCTTGACCG 10080
         •         •         •         •         •         •         •         •
       SalI (10088)                                       KpnI (10139)
CGGTGAGGTCGACTCCGCACCGGGTCAGGGCCTGCTCCGCCCTCGCGCGTATCTCGTCGGTACCACTGATTGCTGCTGTG 10160
         •         •         •         •         •         •         •         •
                StyI (10186)
ACCATgacggtgtgtgctctCTATGCCTTGGATGTGCGGCGGATGTTCACGTAGTTGCCCCCGATCCAATCCTGGAACGT 10240
 <--pcd  •         •<--cefE  •         •         •         •         •         •

GGCGGTCTCGCCGTCCAGGCTGACATCGAAGCCGCACTCGCGCGCCAGCGGGACGGAGAAGGTGAAGTCCGCGTTGGGAC 10320
         •         •         •         •         •         •         •         •

GGAGGAAGAACACACTGGAGGTGCGGCTGCTGCCCGCTATCTGGTCCCTGCGGGGGGCCGCGACATGGTGCCGGGGGGCC 10400
         •         •         •         •         •         •         •         •
                                 NarI (10433)
TTGACCTGGCCGCCGGTCACCAGGGTCGCGATGGCGCCGCAGAAGACGAGGACGGCGTCCGGACGGTAGGGCAGGTCCGT 10480
         •         •    NruI (10426) •         •         •         •         •

GAACGCGCCGCCGACCTCGGCCTGGAGGCTGACGAAGCCGTTGGCGCAGGGTGTCTGCTGGATGAGGGTGACCATCGACA 10560
         •         •         •         •         •         •         •     B5  •
```

```
           NarI (10572)
GGTCGTAGTGCGGCGCCATCCGCAGGGGCTGCTCCTCGGCGCTGCGGTGCTCGGGGACCTGCGGGAAGTAGCGGAACCGC 10640
 B5      .         .         .         .         .         .         .         .
                      StuI (10663)                                    SmaI (10711)
AGCAGCGGCTCGCAGTCGAGGAAGGCCTCGACCCCGCCGTCGGGCTCGGTCCCGGTCGCCCGCAGGACCTCCCGGGCGAC 10720
         .         .         .         .         .         .         .         .
                             ScaI (10750)                           (10798)  SacII
CGCGCGGGAGGCGGTGTACTGGCGGTCGAAGTACTGGGTCCAGATCCGCTCGAAGTCACCGGACGGGAAGAGGTTGTCCG 10800
         .         .         .         .         .         .         .         .

CGGTGCCCATCGAGTAGCACATCGAGTAGTCGGAGTAGCTGCCGGTATTGGTGATCTGGGCGGTGCTCTCCGACTCCAGC 10880
         .         .         .         .         .         .         .         .
          SacII (10890)
CCGGTGAAGCCGCGGCGCATGGTGGGGACGGGCGAGGTGACGGCGCGCTTCTCCGCCTCGCTGCCGTGCTCGAAGAAGTC 10960
         .         .         .         .         .         .         .         .
           StyI (10972)
GATGACGAGGTCCTTGGCCGACTTCAGCTCGGTGTCGGTCAGACCGCAGTCCGTCAGATAGAAGAGGCCCTTGTCCCTCA 11040
         .         .         .         .         .         .         .         .

GACACCTGCGGAACTCGTCCTGGTGCAGGCCCTGCTGGAGTTCGGCCAGGCTGAAGGTGGGCACCGTCGTGTCCATcaac 11120
         .         .         .         .         .         .        <--cefE    .

tctcactgatcctctcgtaatccgtgaaccgggtggtcaacagcgcgggggaggcggtccgcgagccgggggggaTCACTC 11200
         .         .         .         .         .         .         .    <--cefD

GCCCGTCAGCGCGTCCAGCGCGTCCGCCAGGCGGTCGATCTCCTCTTCGGTGGTGTAGAAGTTCGCCGAGATCCGGAGCA 11280
         .         .         .         .         .         .         .         .
    SmaI (11285)        SacII (11305)     Eco47III (11323)
CCGGCCCGGGCGGCTGCTCGGCCACCGCGGCCTCGATCCGGAAGCGCTCCCAGAGCCCCCGGCGCAGCTCGGCCGCGTCG 11360
         .         .         .         .         .         .         .         .
   SmaI (11364)                                                         SmaI (11433)
GTTCCCGGGGGCAGCCGGTAGGCCACCATGCCGCCGGACAGCTCCGGGGAGTCGGGGGTGAGGAGCGTGCGGCCCGGGCG 11440
         .         .         .         .         .         .         .         .
                                                         SmaI (11498)
GTCGGCGAGCAGGCGGCGCGCGTGGTCCGTCAGCTCGCGGCGGCGGGCGCGGATCGCCCCGGGGCCCAGCTCGGCCTGGA 11520
         .         .         .         .         .         . ApaI (11502)      .
   PvuI (11524)                        SmaI (11560)                   SmaI (11591)
AGTCGATCGACTCCGGTGTGGCGAGCCAGGGGCAGATGTCCCGGGTCCCCTCGCATTCGAGTCTGCGCAGCCCGGGTGTG 11600
         .         .         .         .         .         .         .         .

CTGCCGAAGCGGTCGCGCGCGGCCGGCGGGCCGCTGCCCTCCGGGGGCTCGTACGCCCAGCTCACCTGGGTGGGCTCCAG 11680
         .         .         .         .         .         .         .         .

TTCTTCCAGGCGGCCGGGGGCGAGGTGGAGAAAGCCGACCCCGGTGGGGGCGAGCAGCCATTTGTGGCCGCTGCCCGCGT 11760
         .         .         .         .         .         .         .         .
               BamHI (11773)                             NarI (11813)
AGAAGTCGCAGGGGATCCGGGAGAGGTCGAGGTCGAGGAAGCCGGGTGCGTGGGCGCCGTCGACCACGGTGGTGATGCCC 11840
      <--B5      B2-->       .         .         .        SalI (11819)         .

CGTGCGCGGGCCTCCTCGCACAGTTCGGCGGCGGGCAGGATCAGCCCGGTCGTGGAGACGACATGGCTGAAGAAGAACAG 11920
         .         .         .         .         .         .         .         .
                                SacII (11953)
CCGGGTGCGCGGTCCCATCGCGGCGACCGCCGCCGCGGTGATCTCGGCCGGGTCCGAGGGCAGCTCGGGCAGCCGGAACG 12000
         .         .         .         .         .         .         .         .

TCCTCAGCTCCAGGCCGTGCCGCCGGGCGACCCGCTCCCAGCACCAGCGCATGGGCGTGTACTCGTCGTCGCTGAGCAGG 12080
         .         .         .         .         .         .         .         .

ATCTCGCCGGGCGCTTCGAGGCGCAGTGACGACGCGACGAGGTTGACGGCGGCGGTGACGTTGGTGGCCAGGGCGAGGCG 12160
         .         .         .         .         .         .         .         .

CAGCGGGTCCCCGCCGATGAGGCGGGCGAGGCTCTCCCGCGCCTGCCACAGCAGTGCGGGGACCTCGCGGAGCAGGAAGT 12240
         .         .         .         .         .         .         .         .

CCATCGGCTCGGCCGCGAGATGGGCGCGGAAGCCGGTGACCCGCTCGAAGGCGCTGCGCGGCAGCGGTCCCCCGGAGCCG 12320
         .         .         .         .         .         .         .      B2 .
```

```
                                       SacII (12360)
GTGTTGAGGTTGACGACGGTGGGGTCGAGCAGCATACGGCCGCGGGCTTCTTCCCAGTCGGCTACCGCCATgtcaacgca 12400
 B2      .         .         .         .         .         .  <--cefD          .
                                                 PvuI (12450)
tctcccgcaaatccagaagaacagtcaattcagtgcagccatcgaaattcgatcggcccgatgcagcgtagcagtcatga 12480
         .         .         .         .         .         .         .         .

ctcgatctcgaacaggaagattttctgtccgcttcggaaaaatagcgaagagattctgaagaccagccgaagagcacttc 12560
         .         .         .         .         .         .         .         .

aagagtcctgtcgtccagcataggtagcccacattatgcagtaaaccgtccgggctgctcggcatttgaatctttcgacc 12640
         .         .         .         .         .         .         .         .
                            NcoI/StyI (12669)
ttttgtcccgctccacaggagttatcacccatgggtgcacaacggcgccgcccacgtctccgagctgctaaagcagcagg 12720
         .         .         .         .  NarI (12684)      .         .         .
                                                 ClaI (12770)           ClaI (12793)
tcagagcgggctaggcgtgcagattaggccatttcatgactcattttccatcgatctcaacttctggtcgagatcgatgg 12800
         .         .         .         .         .         .         .         .
         SphI (12810)                      BclI (12845)
ttgcgtgatgcatgcggcacctggaccacatcttgtaaaggccgtgatcaccgggtcttgacagttccttcaccggcctg 12880
         .         .         .         .         .         .         .         .
                  NarI (12899)                           PvuI (12938)
gagtgtactgtccagagcggcgcctgacactggttcatacctctgtcgcccgaccgccgatcgaccgccggcggaccatc 12960
         .         .         .         .         .         .         .         .
 Cla I (12962)PvuI (12975)      BclI (12993)
aatcgatggtcgtccgatcgccgcacgatcactgatcaagccctgccccaccccggcgcaaaaattatccgccgggtcat 13040
         .         .         .         .         .         .         .         .

cacgcgtcatcaccacgtaattcgggacgggaataccctgccgcattcccgccggggcggcccggtggggcggctcccct 13120
         .         .         .         .         .         .         .         .

cccgcaggcgcggcggtgcccgtcgcacactcaaggaagagggcaATGAACGACTACTCTCGTCAGAACTTCCTGGATCT 13200
         .         .         .         .     cmcI-->       .         .         .

GAACCTGTTCCGGGGTCTGGGTGAGGACCCGGCCTATCACCCTCCCGTCCTGACGGACCGGCCGCGTGACTGGCCCCTCG 13280
         .         .         .         .         .         .         .         .

ACCGCTGGGCCGAGGCCCCGCGCGACCTCGGCTACTCGGACTTCTCCCCGTACCAGTGGCGCGGGCTGCGGATGCTCAAG 13360
         .         .         .         .         .         .         .         .

GACCCCGACACCCAGGCCGTCTACCACGACATGCTCTGGGAGCTGCGCCCGCGCACCATCGTGGAGCTGGGTGTGTACAA 13440
         .         .         .         .         .         .         .         .

CGGCGGCTCGCTGGCCTGGTTCCGCGATCTCACCAAGATCATGGGGATCGACTGCCAGGTCATCGGCATCGACCGGGACC 13520
         .         .         .         .         .         .         .         .

TGTCCCGATGTCAGATCCCCGCGTCCGACATGGAGAACATCACCCTGCACCAGGGCGACTGCTCCGACCTCACGACCTTC 13600
         .         .         .         .         .         .         .         .

GAGCACCTGCGGGAGATGGCCCACCCGCTGATCTTCATCGACGACGCCCACGCCAACACCTTCAACATCATGAAGTGGGC 13680
         .         .         .         .         .         .         .         .
 SalI (13682)                                              KpnI (13740)
GGTCGACCATCTCCTCGAAGAGGGCGACTACTTCATCATCGAGGACATGATCCCGTACTGGTACCGGTACGCCCCGCAGC 13760
         .         .         .         .         .         .         .         .
                 NarI (13778)          NcoI/StyI (13800)
TTCTCTCCGAGTACCTCGGCGCCTTCCGCGACGTCCTGTCCATGGACATGCTGTACGCGAACGCCAGCTCGCAGCTCGAC 13840
         .         .         .         .         .         .         .         .
SacII (13840)
CGCGGCGTCCTCCGCCGAGTGGCGGCCAAGCAGTAGccaccccaggcccggcgcgtccgtccacggctcagggaagaggt 13920
         .         .        cmcI-->  .         .         .         .         .

ggccATGCCGACCGTATCGAGTCTCCTGTACTACAGCGCACCGCTCACGCCGGGAGGACGGGACGACGACTGGTGCATCG 14000
     cmcJ-->       .         .         .         .         .         .         .

ACGCCGTCTCCCGGCCGCCCGAGGCGCTCTTCAACTTCCGCAAGGTCCCGCACCGGACCGCCGTCACCGATGTGCGCGAC 14080
         .         .         .         .         .         .         .     B2 .
```

```
                                                         SmaI (14136)   ApaI (14151)
GGCACCGTCTCCCCGACCCTCGACGGGATGGCTTCGAGATGGCCGTCGCGCCCACCCGGGCGGACCAACGGGCCCTGCT 14160
 B2      •         •         •         •         •         •         •         •
         NarI (14170)    KpnI (14187)
GGAGGGCTCGGCGCCGGCCGTCGAGCGGTACCAGCGGGAGACCGAAGAGCTGCTGACCGGGCTGACCGGTGCCGACGCGG 14240
         •         •         •         •         •         •         •         •

TGGAGTTCTTCGACGTCACCCTGCGGCGGGAACATGGGGACGCCCCCGCCGACCCGGCCCATCAGTCGCCGCATCTGCGG 14320
         •         •         •         •         •         •         •         •
     SalI (14327)SacII (14339)                SacII (14370)
GTGCACGTCGACCAGAGCCCGCGGAGCGCCTGGGCCAGGGCCGTGGACACCGCGGGGCCCGGCCGCCGCTTCAGCCGGTT 14400
         •         •         •         •         •    ApaI (14375)  •         •
                 NarI (14418)
CCAGATCATCAATGTCTGGCGCCCGCTGATCGAGCCCGTCCGCAACTTCCCCCTCGCGCTGTGCGACTACCGGTCCCTGG 14480
         •         •         •         •         •         •         •         •

ACCCCGCCGCCGATCTGGTGGTCACCCGGCTGGACTTCCCCGCCTGGCTCAAGGACCGCGAGAACTACTCCGTGCGGCAC 14560
         •         •         •         •         •         •         •         •
                KpnI (14577)
AACCCCGGCCACCGCTGGTACCACTGGGGGGCGCTCTCGCCCGACGAGGCGCTGATCTTCAAGTGCTACGACAGCGCGAG 14640
         •         •         •         •         •         •         •         •

CCTCGGTCTGGCCCGCACGGGCCACGGCCCCGAGGACGACGGCACCGCCGGCGACGGCACGGACGCCGGCGGGCGGGCGG 14720
         •         •         •         •         •         •         •         •
         SacII (14729)    SalI (14747)                                    ApaI (14795)
CAGCAGGGCCGCGGCGCGGCGGCCTCGTCGACACCGCCGGGCTCAGTCCGCACACGGCCTTCTTCGACGAGGAGGGGCCC 14800
         •         •         •         •         •         •         •         •

ACCACCGGACATCTGCGCATTTCTCTGGAAATGCGCGCTCTTATCTTTCACGACTGAcgaccgaacgtcgatccggcagc 14880
         •         •         •         •          cmcJ-->  •         •         •
                                        NarI (14921)
cgagaacacccccgacggccgacgccgaggggccgtcgggggcgccaggacagcttcgaaggagcaggggaacATGGCGG 14960
         •         •         •         •         •         •         •     cefR-->
                                         NarI (15002)
ACACGCCCGTACCGATCTTCAACCTCGCCGCACTGCGGGAAGGCGCCGATCAGGAGAAGTTCCGCGAGTGCGTGACCGGG 15040
         •         •         •         •         •         •         •         •
                           NarI (15068)
ATGGGGGTCTTCTACCTCACCGGGTACGGCGCCGGGGATAAGGACCACCGGCTGGCCACGGACACGGCGATGGACTTCTT 15120
         •         •         •         •         •         •         •         •
NruI (15120)
CGCGAACGGCACCGAGGCCGAGAAGGCGGCCGTGACCACGGACGTCCCGACCATGCGGCGCGGCTACTCCGCGCTGGAGG 15200
         •         •         •         •         •         •         •         •
                                                            NcoI/StyI (15261)
CCGAGAGCACCGCCCAGGTGACCAGGACCGGTTCCTACACGGACTACTCGATGTCCTTCTCCATGGGCATCTCGGGCAAC 15280
         •         •         •         •         •         •         •         •
                                     ScaI (15318)
GTCTTCCCCTCGCCGGAGTTCGAGCGGGTGTGGACGGAGTACTTCGACAAGCTCTACGCGGCGGCCCAGGAGACGGCGCG 15360
         •         •         •         •         •         •         •         •

GCTGGTGCTGACCGCGAGCGGCGGCTATGACGCGGAGATCGTCGGAAGCCTGGACGAGCTGCTGGACGCCGACCCCGTGC 15440
         •         •         •         •         •         •         •         •

TGCGGCTGCGGTACTTCCCCGAGGTGCCCGAGCACCGGTCCGCCGAGCACGAGCCGCGCCGGATGGCCCCGCACTACGAC 15520
         •         •         •         •         •         •         •         •

CTGTCGATCATCACCTTCATCCACCAGACGCCGTGCGCCAACGGCTTCGTCAGCCTCCAGGCCGAGATCGGCGGCGAACT 15600
         •         •         •         •         •         •         •         •

GGTGAGCCTGCCCGTCGTGGAGGACGCCGTCGTCGTGATGTGCGGCGCGATGGCCCCGCTGGCGACCCAGGGCGCGCTGC 15680
         •         •         •         •         •         •         •         •
                              NarI (15710)
CCGCGCCCCGGCACCACGTCCGGTCCCCCGGCGCCGGTATGCGCGAGGGCAGCGACCGCACGTCGAGCGTCTTCTTCCTG 15760
         •         •         •         •         •         •         •         •
                                StyI (15792)              SalI (15818)
CGCCCCACGACCGACTTCTCGTTCTCGGTGGCCAAGGCCCGCAGCTACGGCCTCGCCGTCGACCTCGACATGGAGACGGC 15840
         •         •         •         •         •         •         •       B2•
```

```
CACCTTCGGCGACTGGATCGGCACCAACTACGTCACCATGCACGCGAAGAACGAGCCGCAGGCCGGATGAgcagtcctga 15920
 B2      .         .         .         .         .         .  cefF-->          .

caccacggtggggatcggaacacggacggccgccgacggcggccgaaggcgctggaaggacagacATGCTCGTCGTTGCA 16000
         .         .         .         .         .         .     cmcH-->       .
                              PvuI (16032)
TTCAAGCCGGGACATGACGGAGCGGTCGCCGCGATCGACGACCGCCGGCTGCTCTACTCGCTGGAGTCGGAGAAGGACTC 16080
         .         .         .         .         .         .         .         .

CCGGCCGCGCTACTCGACGCTGCTGCCGACGACCTTCCTCGACATCGCCGAACGCCTCGGCGCGATCCCGGACGTCGTGG 16160
         .         .         .         .         .         .         .         .
                                               NarI (16208)
CGCTCGGCGGCTGGGCCGACCTGCGTCCGCGCGGTGTCGTGTACACGGGCGCCGGATACGAGGGCACCCAGGAACCCACC 16240
         .         .         .         .         .         .         .         .
                                                       Eco47III (16296)
GTCACCACCTCGCGCTTCTTCGGCAAGGAGGTGAAGTTCTTCACCTCCACGCACGAGCGCTCGCACATCTACATGGCGCT 16320
         .         .         .         .         .         .         .         .
                   NarI (16340)                                    (16397)   SalI
GGGCATGGCCCCGAAGGACGGCGCCCCGCTGAAGAGCGTGCTCGTCTGGGAGGGGGACGTGGGCGCGTTCTACCTCGTCG 16400
         .         .         .         .         .         .         .         .
                                            SmaI (16445)
ACTCCGAGCACCGGATCATCCGGACGATTCCGGTGATGACGGGACCCGGGGCGCGCTACTCCTTCCTCTTCGGGCTCGCC 16480
         .         .         .         .         .         .         .         .
                              SacII (16511)
GACCCGACCTTCCCGGACACCGGCGGCAAGCCGCGGCTGAACGACGCGGGCAAGCTGATGGCCCTGGCGGCGTTCGGCGA 16560
         .         .         .         .         .         .         .         .
                                                                     NarI (16630)
CTCGGCCGACGCGAGCCCCGACATCCGGCATGTCGTCGAGCGCGTCCTGAAGCAGGACTCCATGTATCCGGCGCCGAAGG 16640
         .         .         .         .         .         .         .         .

CCGAGTACCGGGACTCCGTGCTGCACAACGCGGGCGTCGAGTCGGAGGAGTGCAAGATCGCCGCCGCGTTGCTGACCGAG 16720
         .         .         .         .         .         .         .         .

CGGCTCTTCGAGACCTTCGCGGAGGTCGCCCGGCGCGAACTGCCCGAGGGCACCCCCCTCTACATCTCCGGCGGCTGCGG 16800
         .         .         .         .         .         .         .         .
                          ApaI (16827)                      NarI (16861)
GCTGAACTGCGACTGGAACAGCCAGTGGGCCCAGCTCGGCCACTTCTCCTCGGTCTTCGTGGCGCCCTGCACCAATGACT 16880
         .         .         .         .         .         .         .         .
                     PvuI (16902)
CGGGCTCGGCGCTGGGCACCGCGATCGACGCGCTCACCACGTTCACGGGCGACCCCCATATCGACTGGAACGTCTACAGC 16960
         .         .         .         .         .         .         .         .

GGTCTGGAGTTCGTCCACGACACCCGGCCCGACCCGGCCCGCTGGGAATCCGTCCCGCTCGACCACGCCGCGCTGTCCTC 17040
         .         .         .         .         .         .         .         .

GGCGCTCGCGGCCGGGCGGGTCGTCGCCTGGGTGCAGGGCCGCTGGGAGATCGGCCCGCGCGCGCTGGGCAACCGGTCCC 17120
         .         .         .         .         .         .         .         .

TGCTGGCCGAGGCGTTCAGCTCGGCCAGCCGCGACCGGCTCAACACCGTCAAGATGCGGGAGGACTACCGGCCCATCGCC 17200
         .         .         .         .         .         .         .         .

CCGGTCTGCCGGGTCGAGGACCTGGGGAAGGTCTTCCACGAGGACTTCGAGGACCCGCACATGCTGTACTTCCGGCGGGT 17280
         .         .         .         .         .         .         .         .

CCGGGAGGAGAGCGGGGTACGGGCCGTGACCCATGTGGACGGCTCGGCGCGCTGTCAGACCGTCACCGGGACGAGCAATC 17360
         .         .         .         .         .         .         .         .

CCGAGCTGCACCGGCTGCTGTCGGTGTTCGCCCAGGGGCAGGGCCTCGGTGTCCTGTGCAACACCTCGCTGAACTTCAAC 17440
         .         .         .         .         .         .         .         .

GGCGACGGGTTCATCAACCGGATGTCCGACCTCGTTCGTTACTGCGAATGGCGCGAAATTCAGGACATGGTCGTCGGGGA 17520
         .         .         .         .         .         .         .         .
      KpnI (17526)                       SmaI (17561)SmaI (17572)  SmaI (17587)
CACCTGGTACCGGCGAATAGCCGGTTCCTGAgccgtccgccccgggccgcgcccgggaaaaccgaccccggggtacgttt 17600
         .         .  cmcH-->          .         .         .         .      B2 .
```

```
                           SmaI (17628)               NarI (17655)       SacII (17674)
ttccctcgcaccgagggaagcacggaccccggggtcccggccgaacgggagaggggcgcccggagcggcccgcccgcggg 17680
 B2      .         .         .         .         .         .         .         .
               NdeI (17696)                  NdeI (17726)
tccgcccgtttcgaacatatgcacgtgggagaggtgaccgatacgcatatgccgacgacatgaacgatttccggtcgccc 17760
         .         .         .         .         .         .         .         .
       NarI (17768)
gatccccggcgccgggaaaccctccgagacccttctccgccaccgggaaaacgggaaagcggcggaccggcgacatcggt 17840
         .         .         .         .         .         .         .         .

gcgcaccctccgcgaagaagggcggacaccaccgcagaacaccggaaagcaacgcccttcagacggggaacgaacggggg 17920
         .         .         .         .         .         .         .         .

aaccacagatgacaacgccccccttgtccagaccaattctcttcctgctccaggcatttcatccccgtcaccgaaggcatc 18000
         .         .         .         .         .         .         .         .
                                             SmaI (18046)
gcacacacgtatctccgcgtgattgacggcagcttttatgggcttcccgggaatccggccaactgattgcagattccacc 18080
         .         .         .         .         .         .         .         .
                                                                     SacII (18150)
attccggaggtttgccgaggatttccggactgtcgcccgccgagccccccgtccccaccccggcgcgggccgcggcttcc 18160
         .         .         .         .         .         .         .         .
           BamHI (18172)                      SalI (18220)
gcttccccatcggatccgcccaggtccgggggtagggaggggagagtccgacagccccgtcgacgtcccttcccacagc 18240
     <--B2      B1-->        .         .         .         .         .         .
                                  SmaI (18275)
cttcccacccacccgtcccgactcgccgtgaagccccgggttcttccgggttcaccgaggctgtcccaaatcgtccatgc 18320
         .         .         .         .         .         .         .         .
                                 StyI (18354)                       (18397) StyI
cttgagggtcccgctgcgtgatcgaaccgtaacccttggaatttctgtggattaagcgtaaacatgggtgccgacaccaa 18400
         .         .         .         .         .         .         .         .

ggattacgccgaagccatgtccacccctctcggcgagggcgtggttccttcacaaggggggaccgccATGAACACCTGGAA 18480
         .         .         .         .         .         .     dclX-->       .
                                                               SacII (18544)
TGATGTGACGATCCGGCTCCTGGGGCCGGTGACACTCGTGAAAGGTTCCGTACCGATACCCATCCGCGGGCAGCGACAGC 18560
    ccaR-->        .         .         .         .         .         .         .
                   Eco47III (18580)          StyI (18606)
GGCGATTCCTCGCCTCATTAGCGCTGCGACCGGGCCAGGTCATCTCCAAGGAAGCGATCATCGAAGACTCCTGGGACGGG 18640
         .         .         .         .         .         .         .         .

GAGCCACCACTGACCGTTTCGGGCCAGTTGCAGACGTCGGCCTGGATGATCCGGACCGCGCTGGCGGAGGCGGGGCTGCC 18720
         .         .         .         .         .         .         .         .
                         SacII (18745)
CCGCGACGCCCTCGGCTCCCACGACCGCGGCTACGAACTGCGCGTCCTGCCGGACTCCATCGACCTCTTCGTCTTCCGGG 18800
         .         .         .         .         .         .         .         .

AGGCCGTGCGCGCCGTGCGGGACCTGCACGCACGCGGTCAGCACCAGGAGGCGTCCGAACGGCTCGACACGGCGCTCGCC 18880
         .         .         .         .         .         .         .         .
         ApaI (18890)
CTGTGGAAGGGGCCCGCCTTCGCGGATGTGACCTCCAGTCGGCTGCGGCTGCGGGGCGAGACCCTGGAGGAGGAGCGGAC 18960
         .         .         .         .         .         .         .         .
   SacII (18963)          ClaI (18986)                                   SacI (19034)
CGCCGCGGTCGAGCTGCGCGCCCTGATCGATGTCGGCCTCGGCTACTACGGGGACGCGATCACCCGGCTGTCGGAGCTCG 19040
         .         .         .         .         .         .         .         .
                                             StuI (19086)
TCGATCACGACCCGTTCCGTGAGGACCTGTATGTGAGCCTGATGAAGGCCTACTACGCGGAGGGCCGCCAGGCCGACGCG 19120
         .         .         .         .         .         .         .         .

ATCCAGGTCTTCCACCGCGCGAAGGACATCCTGCGGGAGCAGATCGGCATCAGCCCCGGCGAGCGGATGACAAGGGTCAT 19200
         .         .         .         .         .         .         .         .
                                        KpnI (19241)
GCAGGCCATCCTGCGTCAGGACGAGCAGGTCCTGCGGGTCGGTACCCCGGCCTGAaaccgcgcgcgatacgggaatgttt 19280
         .         .         .         .         ccaR-->   .         .         .
SalI (19281)
gtcgacgtttccctgaaccaacgctgaagaaacgttcttcttctcacaacggcggggaatctccggtcgggaggtgaccg 19360
         .         .         .         .         .         .         .      B1 .
```

```
                                    SmaI (19397)
gaggagcctgcgaagacgtacgacccaccggccccacccgggccggccgggccgagtgccgcccgccgtgcgctccggct 19440
 B1      .         .         .         .         .         .         .         .

ctccttttcttcctcttcttcctcacacagagcagaccgaaaaccaccgccgcgttcacccgccggccgtaacgtccctc 19520
         .         .         .         .         .         .         .         .
             NruI (19535)
gctcacagcgatcatcgcgacagcgcgctcaccgagcgcggggagcaggggcgacagtggccgggcggcggcggacggta 19600
         .         .         .         .         .         .         .         .
      ApaI (19608)                                           NarI (19667)
ctcaaaggggccctggcggcatccggtgcgccggcggtcggggcagggctcggtgtcggcctcgccggcgccgacgcgcc 19680
         .         .         .         .         .         .         .         .
     SacII (19687)      SmaI (19707)
cgcccgccgcgggggctccttctccacccgggtgcggctgcgctggctgggggtctccggctgggagatcgtcatcgacg 19760
         .         .         .         .         .         .         .         .
                                                    SacII (19814)
gcgggcacagcattctgtttgacccgtatctgagccggATGCCCTGCCGGAATCCGCGGACGGGCGCGCTCGATTCACGG 19840
         .         .         .      orf11--> .         .         .         .
                                                                   (19917)  BclI
CTGCCCCTGCGGACCGACCGGTCCCTGGTGGAGGCGACGGCCGCGCGCCATCTCACCGTTCCGCCCGAGTTGATCCTGAT 19920
         .         .         .         .         .         .         .         .
   NcoI/StyI (19924)
CAGCCATGGCCATTTCGACCATATGGCCGATGTCCCGCAGTTGTTGGCCCGTCCGGCGTGGCGGCGGCGGCGCATCCGGA 20000
         .         NdeI (19940)        .         .         .         .         .

CCGTGTGCGACGAGACGGTGCTGCATCTGCTGCTGGCGCTCGACGCGCCCCGGCCCCGGATCGACGACGTCATCCAGGTG 20080
         .         .         .         .         .         .         .         .

AAGGGCGGCGAGTATCTGCGCTTCGACGGCTACACGGTGGAGGTCTTCCGCAGTCTGCACAGCCGGCTGCCGGACCACGG 20160
         .         .         .         .         .         .         .         .
         SmaI (20170)         SacII (20192)
CTCCTTCGCCCCGGGGCGGCGGGTGTCCCCGCCGCGGCGGCCGTCCGCCCTGGGCGATCTGGTCGAGGGCGACACCCTCG 20240
         .         .         .         .         .         .         .         .
               SalI (20255)
CCTACCAGGTCTCCGTCGACGGCGGGCCCACGATCCTGCTGATGGGCGCGGGGAACTTCGTCGAACGCGAACTGGCCGGA 20320
         .         .   ApaI (20264)    .         .         .         .         .

ATCCGGCCCGATGTCGCGCTCGTCCCCATGACGTCGCACACGGCGGTGCACCGCTATGTCGAGCGGCTGCTGACGGCAAC 20400
         .         .         .         .         .         .         .         .

GGGAAATCCGCCGCTGCTGATCCCCAGTCACCACGACGACATGACCACGCCGCTCACCCGTGCGCCTCGGGCGCTCGCCG 20480
         .         .         .         .         .         .         .         .
                                 SmaI (20514)
ACTCGGTCCGGCCGATGGCGGCACGGGAGCTGTCCCGGGCCGCCGCAGCGCGATCACGCGTGATCGAACCGCATCACCTG 20560
         .         .         .         .         .         .         .         .

GGTGTGCTGGATCTCACCGCCGCACTGTCCTGAcctgcgcgaaagccctgaatgaccgatgcgggccagggtctgtcacc 20640
         .         .     orf11-->      .         .         .         .         .

agccgttggaaggcggcggagatcagcggaagaccacggctggtgaggatggcccgcacaccggttcttgcacctcacgc 20720
         .         .         .         .         .         .         .         .

accttccgcacctcgtgcactttcgcacctcgtacgtatcccgcgccttgcccatatcccgcacctcgtacgtatcccgc 20800
         .         .         .         .         .         .         .         .

gccttgtacgtatcccgtatctcgtgcgccttccgcacctcgtgcacctctcacaccccatgcatccccagcacgtcatg 20880
         .         .         .         .         .         .         .         .

cacctcccggacctcatgggccgacccagacagccgatacagacaaggggttccttcacATGGTGAAGAAGACATGGAGG 20960
         .         .         .         .         .         blp-->    .         .
                   SacII (20977)
TCCGCGCTCGCGCTGACCGCGGCCGCCGGAGCCGTCCTCGCCACGGCCACCAGCGCCCACGCGTACACGGGCTTCACGCC 21040
         .        NotI (20979)         .         .         .         .         .

CGAGCGGTACAACAAGATCCAGTTCGGGATGGACCGGACCCTGGTCTGGCAGCTCGCCGGTGCCGACCAGTCCTGCTCGG 21120
         .         .         .         .         .         .         .     B1 .
```

```
                                                                   SacII (21193)
ACCAAGTGGAGCGGATCATCTGCTACAACAACCCGGACCACTACGGGCCGCAGGGCCACTTCTTCTTCAACGCCGCGGAC 21200
         .         .         .         .         .         .         .         .
 B1

AAGCTCATCCACAAGCGGCAGATGGAGCTGTTCCCCGCGCCGAAGCCGACCATGCGGCTGGCCACCTACAACAAGACCCA 21280
         .         .         .         .         .         .         .         .

GACGGGCATGACGGAGGCCCAGTTCTGGGCGGCCGTCCCCTCGGACACCTGCTCGGCGCTGGCCGAGCAGTACCCGAACT 21360
         .         .         .         .         .         .         .         .

GGCCCGCCACGAACGGCAATCTGCGGGAGTACGTCTGCCCCAGCAAGGCCGAGCGGTTCGCGCCGAGCGCCTATTTCACC 21440
         .         .         .         .         .         .         .         .

TTCACCGATGGCAAGCTGACCAGCAGGTCGCAGAGCCAACTGCCCTGAagcgggccatagcagaccgaaggccgggcttc 21520
         .         .         .         . blp-->  .         .         .         .

gggagaaacaccgaaaggtgatgttccactgccgggccgacgatgcgccgtttccggaccgaacgcatgattcgggtttt 21600
         .         .         .         .         .         .         .         .
                                                        PvuI (21657) (21677) PvuI
ccaccgcgctgccgaccggaaaacaccggcgcaatcccatcgaacgcacccctgcccgatcggggcgttccacccccgat 21680
         .         .         .         .         .         .         .         .

cgggcagcgcggccaccgaatgtccttcggtggccgttccgcttttccggggcggccgggcagcacccgagtacccggcc 21760
         .         .         .         .         .         .         .         .

gaaaacggtcggtcccaggcttcgatggccaccgcaaagatccagccatgcgccgccctccgcgcctccgccccccacct 21840
         .         .         .         .         .         .         .         .
                          EcoRI (21867)
attcggaaaaggcggcaagcccgttggaattcccctgaacacgaagctgagcaacagctcgtcacgcgctcccgagctgg 21920
         .         .         .         .         .         .         .         .

ccattcagggcagttcacaaagagccatcgagaggcgtccgagagagctggaagaggggtccaagagcatggtgggtcat 22000
         .         .         .         .         .         .         .         .

tattgtgatcctaaaatgtccagttcaccgccatgacagcagaggctggaaagtcccccataattcagcctgatccccca 22080
         .         .         .         .         .         .         .         .
           NcoI/StyI (22092)
ggagttctcacccATGGGCGAAGCAGCACGCCACCCCGACGGCGATTTCTCGGACGTGGAAACCTCCACGCTCAGGACG 22160
         .  lat-->           .         .         .         .         .         .
                    NdeI (22181)
TGCACCAGGCACTTGAGCAGCATATGCTCGTCGACGGGTACGACCTCGTTCTCGACCTCGACGCCAGCTCCGGCGTCTGG 22240
         .         .       SalI (22190)          .         .         .         .
   SalI (22244)                                         NarI (22297)
CTCGTCGACGCCGTCACCCAGAAGCGGTATCTCGACCTCTTTTCCTTCTTTGCCTCGGCGCCGCTCGGAATCAACCCGCC 22320
         .         .         .         .         .         .         .         .
                                             Sac II (22365)
CAGCATTGTCGAGGACCCGGCATTCATGCGGGAGCTGGCCGTGGCCGCGGTCAACAAGCCGTCGAACCCCGATCTTTATT 22400
         .         .         .         .         .         .         .         .
                                 SmaI (22434)
CGGTGCCGTACGCCCGTTTCGTCAAGACCTTCGCCCGGGTCCTCGGCGACCCCCGGCTGCCGCGGCTGTTCTTCGTGGAC 22480
         .         .         .         .         .         .         .         .

GGCGGGGCGCTGGCCGTGGAGAACGCGCTCAAGGCGGCCCTCGACTGGAAGGCCCAGAAGCTGGGCCTCGCCGAGCCGGA 22560
         .         .         .         .         .         .         .         .
                             Eco47III (22590)
CACCGACCGGCTCCAGGTGCTGCATCTGGAGCGCTCGTTCCACGGCCGCAGCGGCTACACCATGTCGCTGACGAACACCG 22640
         .         .         .         .         .         .         .         .

AGCCGTCCAAGACCGCCCGCTTCCCCAAGTTCGGCTGGCCACGGATCTCGTCCCCCGCCCTCCAGCACCCGCCGGCCGAG 22720
         .         .         .         .         .         .         .         .
      NarI (22727)                               SmaI (22767)
CACACCGGCGCCAACCAGGAGGCCGAGCGACGGGCGCTGGAGGCCGCCCGGGAGGCGTTCGCAGCGGCGGACGGCATGAT 22800
         .         .         .         .         .         .         .         .

CGCCTGCTTCATCGCGGAGCCCATCCAGGGCGAGGGCGGCGACAACCACCTCAGCGCGGAGTTCCTCCAGGCCATGCAGC 22880
         .         .         .         .         .         .         .    B1  .
```

```
                                                              KpnI (22946)
GGCTCTGCCACGAGAACGACGCCCTGTTCGTCCTGGACGAGGTGCAGAGCGGCTGCGGCATCACCGGTACCGCCTGGGCC 22960
 B1      .         .         .         .         .         .         .         .

GATCGACGAGGTCCCCGAGAACGTCTTCGCCGTCTCCTCCCGGATCAGCTCCACCTGGGGCGGCAACCTCGCCGACATGG 23120
         .         .         .         .         .         .         .         .
                    PvuI (23145)                                    ScaI (23190)
TCCGCGCCACCCGGCTGCTGGAGACGATCGAGCGCACCCAGGTCTTCGACACCGTCGTCCAGCGCGGCAAGTACTTCCGG 23200
         .         .         .         .         .         .         .         .
                                                       SacII (23251)    (23276)  SalI
GACGGCCTGGAGGACCTGGCCGCCCGCCACCCCTCCGTCGTGACCAACGCCCGCGGCCGGGGCCTGATGTGCGCGGTCGA 23280
         .         .         .         .         .         .         .         .

CCTGCCGGACACCCGGACCCGCAATGAGGTGCTGCGGCTTATGTACACGGAGCACCAGGTCATCGCCCTGCCCTGCGGCG 23360
         .         .         .         .         .         .         .         .
                              PvuI (23391)
GGCGCAGCCTCCGGTTCCGCCCCGCGCTGACGATCGCCGAGCACGAGATCGACCAGGCCCTTCAGGCGCTGGCGAGCAGT 23440
         .         .         .         .         .         .         .         .
                                                          ApaI (23494)     SacII (23512)
GTCACGGCGGTCGCCGAGAGCGTCTGAcgcccggccggccgggtgccccagcggggcccggccggcccgcaccgcggaac 23520
         .         ·lat-->   .         .         .         .         .         .

ggaacacccctggcgcgtcgcccggtgacacgctctccgggcggctcaaccgcgccacccccacggcacgttcgccttca 23600
         .         .         .         .         .         .         .         .
                                 KpnI (23634)
cccgcacggagagcccacgaATGATGTCAGCACGGTACCCGAGGACCGCAGCGGAGTGGACCACTCGCATTCAAGGAGTG 23680
         .         ·pcbAB-->           .         .         .         .         .

TCGAGCGAGCGTTGCGATCTTGAAATGCTGCTGAAGGACGAGTGGCGCAACAGGATCGCGGTACGGGACGACGACCCCGG 23760
         .         .         .         .         .         .         .         .
                               SalI (23792)
TGTCCGTGCGACGAGGCAGCGGGACATCGTCGTCGACGGGCGGGAGTACACCGCTCTGAAGGACGCGCTGCGCGCCGCCG 23840
         .         .         .         .         .         .         .         .
                                                             NcoI/StyI (23902)
ACGGGGTCTCGGCGGGCGCGCTCGCGCTGGCCTCCCTGCACAGCGTGATGCGCGCGTACGGCCATGGCGAGCAGACGGTG 23920
         .         .         .         .         .         .         .         .
          SalI (23930)                                         SalI (23984)
GCGGCGTTCGTCGACGCGACGGCGACGGCGGAGCTGAAGACGGCCGCCGTGCTGCCGGTGATCGTCGACCATATCGAGCA 24000
         .         .         .         .         .         .         .         .

CACCCGGCTGACCTGCGCCGAGGCGATCCGGGAGCTGGACGAGACGCTGCGCCGCAAGGACTCCTACACCCGTGCGGACG 24080
         .         .         .         .         .         .         .         .

AGGTGCTCCAGCGGGGGCTCTTCGATGCCCTGCTGGTGCTCGCCGAGCGCGAGGTGGCCCTGTCCGAGCTGCCGTCGGCC 24160
         .         .         .         .         .         .         .         .
                               SmaI (24192)
CCGCTGGTCATGGTCGTCCGGGACGACGCGGCCCGGGGCCGGCTGTGCTGGACGATGGCGTACGCGGGCGAGCTGTTCGA 24240
         .         .         .         .         .         .         .         .

GGACACGACGGTCGCCGGTGTGCTGGAGGTCGTCCGGGAGGTGCTCGGGCAGTACGCCGGGCGGCCGGGCGACCGGGTCG 24320
         .         .         .         .         .         .         .         .

CCGAGATCGAGCTGGCCTCGCGGGAGCAGCGCGAGCGGCTTCAGCGGTGGAACGCGACCGATGGCGACTTCCCGGCGGAC 24400
         .         .         .         .         .         .         .         .

CAGCGCCTCAACGACCTGGTCGAGGCGGCCGTGCGGCGGTCGCCGGACCGTGAGGCGGTCGTCTTCGGGACACAGCGGCT 24480
         .         .         .         .         .         .         .         .
           SalI  (24493)
GACCTATCGGGAGGTCGACGCGCGGGCGAACCGGTTCGCGCACTGGCTGCTCGGTCCGGGTCTCGGGGTCCGTTCGCAGC 24560
         .         .         .         .         .         .         .         .

AGCTCGTCGGGATCTTTCTGGACAAGAGCGATCTCGGGGTGGTCGCCACGCTCGGCATCTGGAAGGCGGGCGCGGCCTAT 24640
         .         .         .         .         .         .         .         .
    PvuI (24645)
GTGCCGATCGACCCGGCCTATCCGGCGGAGCGGGTGCGGTTCGCGGTCGGTGACACCGGGCTGCGGGGCATCGTCACCAA 24720
         .         .         .         .         .         .         .    B1 .
```

```
CCGGCACCACGCGGGGCGGCTGCGGGAGATCCTGGGCGCGGAGCACGCGGATGTGACGGTCGTCGAGATCGAGTCCGTGC 24800
 B1      .         .         .         .         .         .         .         .

TCGACGAGCAGGCCGCCGCCGACACCGACGGCCTGCTGAGTGTGAAGCCCGAGCTGGCGCTGGGGGTACGGGATCTGGCC 24880
         .         .         .         .         .         .         .         .

TATCTCACCTACACCTCCGGCACCACCGGTGTGCCGAAGGGGGTGCCCAAGTACCACGACAGTGTCGTGAACAGCATCAC 24960
         .         .         .         .         .         .         .         .

GGACCTGTCGGAGCGGTACGACATGCGGCGGCCGGGCACGGAGCGGGTCGCGCTCTTCGCCTCGTACGTCTTCGAGCCGC 25040
         .         .         .         .         .         .         .         .

ATCTGCGGCAGACGCTGATCGCGCTCATCAACGGGCAGACGCTGGTCGTCGTCCCCGAGGAGGTCCGGCTCGACCCGGAC 25120
         .         .         .         .         .         .         .         .

CGCTTCCCCGCGTACATCGAGGAGCACGGCGTCACGTACCTCAACGCGACGGGTTCGGTGCTGCAACACTTCGATCTGCG 25200
         .         .         .         .         .         .         .         .

CCGACGTACCAGCCTGAAACGGCTGTTGCTGGTCGGCGAGGAGCTGACGGCGGCCGGGCTGCGCCAGCTCCGCGAGCGGT 25280
         .         .         .         .         .         .         .         .

TCTCCGGGCGGATCGTCAACGAGTACGCCTTCACCGAGGCGGCGTTCGTCACCGCCGTGAAGAAGTTCGCGCCGGGCGTC 25360
         .         .         .         .         .         .         .         .

ACCGAGCGCGCGGACCGCAGCATCGGCCGACCGGTGCGGAATGTGAAGTGGTATGTCCTCAGCCAGGATCTGAAGCGGTT 25440
         .         .         .         .         .         .         .         .
           PvuI (25452)                      NarI (25486)
GCCGGTCGGGGCGATCGGCGAGCTGTACATCGGCGGCTGCGGGGTGGCGCCGGGCTATCTGAACCGCGACGACCTGACGG 25520
         .         .         .         .         .         .         .         .

CCGAGCGGTTCCTGACCAATCCGTACGCCTCGGAGCAGGACCGGGCGCGCGGGACGAACGCCCGGATCTACCGGACCGGT 25600
         .         .         .         .         .         .         .         .

GACCTGGCGCGGATGCTGCCGTCCGGCGAGGTCGAGTTCATGGGCCGCAGCGACTTCCAGCTCAAGCTGAACGGTGTCCG 25680
         .         .         .         .         .         .         .         .

GGTGGAGCCGGGCGAGATCGAGGCGCAGGCCACCGAGTATGCGGGGGTCCGCAAGTGCGTGGTCATCGCCCGTGAGGGGG 25760
         .         .         .         .         .         .         .         .
                                              SmaI  (25808)
CGGGCGGCGGCAGCGACCGCCATCTGGTCGGCTACTACCTGACGGAGCCCGGGGCCGGGGTGACCGAGGCGGAGCTGCTG 25840
         .         .         .         .         .         .         .         .

TCCTTCCTGGAGCGGCGGCTCATCCGGATCATGGTCCCGGCGCGGATGGTCCGGCTGGAGAGCATCCCCGTCAATGTCAA 25920
         .         .         .         .         .         .         .         .
       SalI (25928)
CGGCAAGGTCGACTGGCGGGCGCTGCCCGAGGTGGACCTCGCCCGCCCGGACACGGGCGGCGGCGCGGTGGGCTCCACCA 26000
         .         .         .         .         .         .         .         .

CGACCGGGGGTGTCCGGGGTGAGCTGCGCGAGATC  26035
         .         .      pcbAB-->
```

The nucleotide sequence for the *bla* gene was described in Perez-Llarena et al. (173) with the nucleotide accession number of Z54190. The nucleotide sequence for the *cefD* gene was described in Kovacevic et al. (123) with the nucleotide accession number of M32324. The nucleotide sequence for the *cefE* gene was described in Kovacevic et al. (124) with the nucleotide accession number of M24140. The nucleotide sequence for the *cefF* gene was described in Kovacevic and Miller (122) with the nucleotide accession number of M63809. The nucleotide sequence for the *lat* gene and part of the *pcbAB* gene was described in Tobin et al. (215) and Yu et al. (234) with the nucleotide accession numbers of U12015 and M64834. Differences were noted between the two *lat* sequences (234) and the sequence of Yu et al. (235) was picked to be included in the compilation,since it extended the furthest but the 141 bp downstream of the *Eco*RI site was from Tobin et al. (215). The nucleotide sequence for the *dclX* gene and upstream sequence was provided by B. Barton (19) and was described in Walters et al. (221) but does not have a nucleotide accession number. In all cases, attempts were made to minimize the amount of duplicated sequence data.

A graphic restriction map was generated from the sequence data to provide a demonstration of the results from the sequencing project and the location of the newly discovered ORFs (Fig. 9). Seven ORFs not previously described in *S. clavuligerus* were discovered as a result of the sequencing project and five additional undescribed ORFs were discovered but their sequences were published by others during or after completion of the sequencing project. The new ORFs were labeled with the gene designations that were felt to be appropriate based upon homology to known genes from *Streptomyces* or other organisms (see III.5).

III.4 Analysis of the cephamycin cluster sequence information

Duplicated sequence generated through overlaps or from sequences published after the completion of the sequencing project, were checked for any differences from published sequence. After completion of the sequencing project Perez-Llarena et al. (174) published the sequence of the *ccaR*, *orf11* and *blp* genes with the nucleotide accession number of Z81324. Differences were noted between our sequence and the newly published sequence and they are summarized in Table 6. It is not known whether these discrepancies were due to allelic differences between the two strains or sequencing errors but all of the differences were located within intergenic regions. Careful re-examination of my raw sequence data and those of Walters et al. (19, 221) has failed to show any regions of ambiguity and so these changes may represent genuine differences between the two separate strains despite their common origin.

Figure 9. Diagrammatic representation of the cephamycin biosynthetic gene cluster as determined from the compiled sequence data. Solid arrows indicate the direction of transcription, circles were placed above genes described during or after completion of the sequencing project and asterisks were placed above the genes not previously described in *S. clavuligerus*. The two fragments shown were contiguous, with *cefD* located adjacent to *cmcI*.

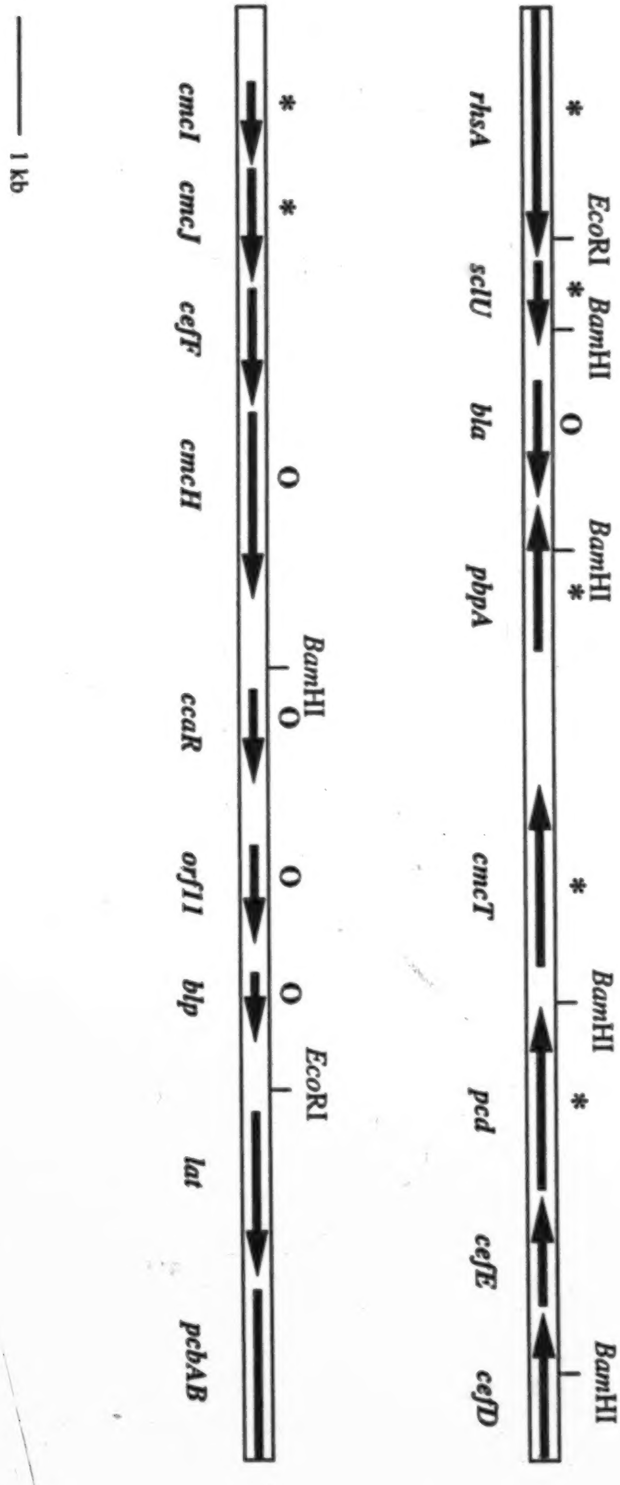

EcoRI
BamHI
BamHI
BamHI
BamHI
rhsA
sclU
bla
pbpA
cmcT
pcd
cefE
cefD
BamHI
EcoRI
cmcI
cmcJ
cefF
cmcH
ccaR
orf11
blp
lat
pcbAB
1 kb

TABLE 6. Differences between published sequence data and new data

sequence location[a]	gene location	published sequence[b]	new sequence
18179-18181	*ccaR* upstream	cc	ccc
18257-18259	*ccaR* upstream	cc	ccc
18351-18352	*ccaR* upstream	a	aa
19403-19404	*orf11* upstream	c	cc
19414	*orf11* upstream	gg	g
19544-19546	*orf11* upstream	ccg	cgc
19630-19631	*orf11* upstream	gg	gc
21655-21657	*blp* downstream	cc	ccc
21662-21665	*blp* downstream	ggg	gggg
21691-21692	*blp* downstream	g	gg
21704-21705	*blp* downstream	c	cc

[a] The sequence location was determined from the compiled sequence data in Figure 8.

[b] The sequence determined in this study was compared to the published sequence of Perez-Llarena et al. (177) determined from wildtype *S. clavuligerus* ATCC 27064.

The sequences also differ in the start codons chosen for the *ccaR* and *orf11* genes. Walters et al. (221) have chosen an ATG start codon for DclX that is located six amino acids ahead of the GTG start codon chosen for CcaR by Perez-Llarena et al. (174). Similarly, Perez-Llarena et al. (174) chose the GTG start codon at nt. 19577 for Orf11 but I have chosen the ATG start codon at nt. 19799, giving a deduced protein of 264 amino acids instead of 338 amino acids. The ATG start codon was chosen because it was believed to be more consistent with the %GC data but again no experimental data are available to indicate which start codon is correct.

The gene described in Xiao et al. (230) was mistakenly referred to as *cmcI* but was not found to be part of the cephamycin cluster and should more correctly be referred to as an ORF from *S. clavuligerus* with hydroxylation activity.

III.5 Characterization of the cephamycin cluster gene products

The majority of the new genes discovered were deduced from the DNA sequence by codon preference analysis and database homology searches. These genes can only be assumed to produce their respective proteins. The cephamycin cluster gene products were characterized with the DNA Strider 1.2 program (144). A hydropathy plot, using the Kyte and Doolittle algorithm (129), the number of amino acids and the calculated molecular weight were all determined from the translated gene product. A summary of the physical properties of the gene products is presented in Table 7. The location of each protein within the cell was predicted using the results of the hydropathy plot analysis but the location of homologous proteins was also considered.

The cephamycin cluster gene products were also characterized by gapped BLASTP (gBLASTP) searches of the databases at the National Center for Biotechnology Information (10). As well, the percentage of identity between the homologous proteins and the *S. clavuligerus* protein was calculated. The results of the database homology searches and the percent identity calculations are summarized in Table 8.

The gBLASTP database search provides two scores for each homologous protein which ranks each homologous protein against the other homologs for its similarity to the query protein (10). The maximal segment pair (MSP) measure gives a matrix of similarity score for all possible pairs of residues and the boundaries of each MSP are chosen to maximize its score (9). The *E*-value is a measure of the probability of finding more distinct MSPs with a score equal to or greater than the score for this MSP. Position-Specific Iterated BLAST (PSI-BLAST) is more sensitive to weak but biologically significant similarities as it searches the databases for motifs found within the query protein (10). PSI-BLAST was used for the proteins which did not display

TABLE 7. Proteins encoded by the *S. clavuligerus* cephamycin cluster

Protein	Number of amino acids[a]	Molecular weight (KD)[a]	Predicted location[b]	Function[c]
RhsA	>704	>74.9	cytoplasmic or extracellular	unknown
SclU	264	26.6	membrane	unknown
Bla	332/305	35.2/32.6	extracellular	β-lactamase
PbpA	409	43.3	membrane	PBP
CmcT	489	49.7	membrane	transport
PCD	512	54.0	cytoplasm	dehydrogenase
DAOCS	312	34.5	cytoplasm	ring expansion
IPNE	398	43.5	cytoplasm	epimerization
CmcI	236	27.6	cytoplasm	hydroxylase
CmcJ	310	34.0	cytoplasm	methyl-transferase
DACS	318	34.6	cytoplasm	hydroxylase
CmcH	522	57.3	cytoplasm	carbamoyl-transferase
CcaR	256	28.6	cytoplasm	transcriptional activator
Orf11	264	29.0	cytoplasm	unknown
Blp	182/154	20.3/17.5	extracellular	unknown
LAT	457	49.9	cytoplasm	aminotransferase
ACVS	unknown	unknown	cytoplasm	peptide synthetase
IPNS	329	36.9	cytoplasm	cyclization
PcbR	551	57.3	membrane	PBP

[a] The number of amino acids and calculated molecular weight of the translated gene product were determined using the DNA Strider 1.2 program.

[b] Predicted location of the protein within the cell was determined by hydropathy plot analysis and the location of homologous proteins in other organisms.

[c] Function of the gene was deduced from earlier characterization in *S. clavuligerus* or *N. lactamdurans* or from homology to proteins characterized in other organisms.

TABLE 8. Proteins homologous to the new *S. clavuligerus* gene products

Protein[a]	Organism[b]	gBLASTP scores[c]	Percent identity[d]	Length[e]	Function[e]
RhsA	herpesvirus	64, 3e-09	24.1 (436)	866	membrane protein
(>704)	*Ba. subtilis*	57, 4e-07	20.5 (453)	2334	unknown
	E.coli, RhsC	55, 1e-06	22.2 (518)	1397	unknown
	E. coli, RhsB	54, 2e-06	21.5 (506)	1411	unknown
Bla	*My. tuberculosis*	273, 9e-73	57.3 (316)	307	β-lactamase
(332)	*N. lactamdurans*	238, 4e-62	52.9 (280)	302	β-lactamase
	Bu. cepacia	227, 8e-59	54.2 (277)	302	β-lactamase
	S. badius	225, 2e-58	50.3 (300)	313	β-lactamase
PbpA	*My. tuberculosis*	69.6, 3e-11	31.2 (208)	291	PBP
(409)	*P. aeruginosa*	68.1, 7e-11	30.0 (240)	311	PBP5
	Ba. subtilis	67.0, 2e-10	29.5 (244)	389	unknown
	E. coli	65.9, 3e-10	30.2 (248)	390	PBP
CmcT	*N. lactamdurans*	395, 1e-109	72.2 (468)	486	efflux
(489)	*S. anulatus*	153, 2e-36	40.3 (477)	503	efflux
	My. tuberculosis	119, 3e-26	36.2 (450)	530	efflux
	S. glaucescens	114, 1e-24	29.4 (418)	538	efflux
PCD	*My. tuberculosis*	515, 1e-145	68.8 (496)	494	dehydrogenase
(512)	*C. elegans*	359, 2e-98	50.9 (485)	514	dehydrogenase
	P. sativum	355, 2e-97	48.2 (479)	508	turgor response
	H. sapiens	343, 2e-97	48.4 (481)	511	dehydrogenase
	Br. napus	327, 8e-89	48.0 (440)	494	turgor response
	E. coli	186, 3e-46	33.3 (474)	482	dehydrogenase
CmcI (326)	*N. lactamdurans*	389, 1e-108	78.3 (235)	236	hydroxylase
CmcJ	*N. lactamdurans*	386, 1e-107	69.6 (263)	288	methyltransferase
(310)	*Me.* sp.	67.4, 9e-11	29.9 (157)	353	unknown
CmcH	*N. lactamdurans*	849, 0.0	82.9 (520)	520	carbamoyltransferase
(522)	*R. tropici*	203, 1e-51	42.7 (253)	575	carbamoyltransferase
	R. sp. NGR234	199, 2e-50	36.9 (293)	558	carbamoyltransferase
	R. fredii	198, 6e-50	39.5 (261)	581	carbamoyltransferase
Blp (182)	*S. clavuligerus*	67.0, 6e-11	31.4 (191)	201	β-lactamase inhibitor

[a] Cephamycin cluster protein and its length in amino acids in brackets.

[b] Sources of the homologous proteins. Abbreviations for the genus names not previously defined are: *Ba.* for *Bacillus*, *Br.* for *Brassica*, *Bu.* for *Burkholderia*, *C.* for *Caenorhabditis*, *H.* for *Homo*, *Me.* for *Methylobacterium*. *My.* for *Mycobacterium*, *R.* for *Rhizobium* and *P.* for *Pisum*.

[c] All homologous proteins were ranked in order with respect to their gBLASTP scores (MSP, *E*-value) calculated during the database searches.

[d] The percent identity of each homologous protein to the *S. clavuligerus* protein was calculated using the LFASTA program over a specific length of amino acids, given in brackets.

[e] Length of the homologous protein in amino acids and the deduced function of the homologous protein.

homology to any known proteins present in the database. If the query protein had already been submitted to the database it would obviously be detected during the database search. In the instances where this happened the *S. clavuligerus* protein was omitted from further analysis.

The goal of database homology searches is to infer the function of a protein by its similarity to a previously characterized protein (39). Unfortunately, many of these homologous proteins have been determined by genome sequencing projects (179) and their actual function has often not been determined but rather was deduced from homology in database searches.

The percentage of identical amino acids can be calculated by two methods both relying of the alignment algorithm of Pearson and Lipman (172). The ALIGN program calculates percent identity over the total protein whereas the LFASTA program calculates percent identity over the smaller homologous region. When using the LFASTA program the size of the homologous region would differ for each protein pair but in all cases the largest window of amino acids compared was chosen and this value was included with the percent identity measure. The ALIGN program gives lower scores for proteins that are much larger or smaller than the query protein because the program extends the region of homology throughout the length of both proteins rather than confining it to the region of strongest homology. These functions give the homologous protein a lower percent identity measure and therefore the values produced by the LFASTA program were used to compare the new *S. clavuligerus* proteins to their homologs in the databases. The rank of the proteins may change over that given by the gBLASTP searches since the percent identity value was not affected by the number of gaps that were inserted into either protein to give maximal alignment.

III.5.1 The *rhsA* gene product

The *rhsA* gene from *S. clavuligerus* has only been partially sequenced but from the sequence data obtained the RhsA protein would be at least 704 amino acids with a calculated molecular weight of at least 74.9 KD. The hydropathy plot of the RhsA protein lacks a concentration of hydrophobic amino acids in one area (data not shown) suggesting that the protein is not membrane associated. Since the N-terminal end of the gene has not been sequenced it could not be determined whether the putative protein was cytoplasmic or extracellular.

The RhsA protein when analyzed by gBLASTP database searches showed homology to a number of *rhs* genes from *E. coli*. The strongest homologs were a membrane glycoprotein from a herpesvirus (120), a wall-associated protein precursor from *B. subtilis* (79, 233) and the *rhsC* (236) and *rhsB* (189) genes from *E. coli*.

The Rhs proteins from *E. coli* are hybrid proteins that consist primarily of a core region that contains a peptide motif that was repeated 28 times and is homologous to a cell surface ligand-binding protein (94). This same repeat was also found in the *B. subtilis* wall-associated protein. These Rhs elements also contain distinct components with %GC compositions which are very different from that of the surrounding genome suggesting that they recently originated outside of *E. coli*. The function of these proteins in *E. coli* is unknown but they are not essential since some strains of *E. coli* do not have any Rhs elements while *E. coli* K12 has five different Rhs elements. The regions are known to be recombination hot spots and for this they are named.

The function of the *rhsA* gene in *S. clavuligerus* is unknown but it seems unlikely that it is involved in cephamycin biosynthesis.

III.5.2 The *sclU* gene product

The *sclU* gene encodes a 264 amino acid protein, SclU, with a calculated molecular weight of 26.6 KD. The hydropathy plot of the SclU protein showed a number of transmembrane domains capable of inserting the protein into the cytoplasmic membrane (data not shown).

The SclU protein does not show significant homology to any known proteins when analyzed by gBLASTP or PSI-BLAST database searches. The gene designation *sclU* was chosen for this gene because it encodes an *S. clavuligerus* unknown protein.

III.5.3 The *bla* gene product

The *bla* gene encodes a 332 amino acid protein, Bla, with a calculated molecular weight of 35.2 KD. The *bla* gene was sequenced and characterized by Perez-Llarena et al. (173), who determined that the Bla protein was present in a processed form associated with the cell wall. The pre-β-lactamase signal peptide was processed during transport from the cell, to give a final 305 amino acid β-lactamase with a molecular weight of 32.6 KD.

The Bla protein showed homology to a very large number (>50) of type A β-lactamase precursors primarily from *Streptomyces* and *Bacillus* when analyzed by gBLASTP database searches. The strongest homologs were type A β-lactamase precursors from *Mycobacterium tuberculosis* (direct submission), *N. lactamdurans* (51), *Burkholderia cepacia* (direct submission) and *S. badius* (78). Perez-Llarena et al. (173)

showed the Bla protein to contain all of the conserved motifs typical of β-lactamases but the purified protein only had weak β-lactamase activity.

The Bla protein may function as a resistance mechanism to provide protection to the cell from the antibiotic that it produces.

III.5.4 The *pbpA* gene product

The *pbpA* gene encodes a 409 amino acid protein, PbpA, with a calculated molecular weight of 43.3 KD. The hydropathy plot of the PbpA protein showed a single stretch of hydrophobic amino acids in the N-terminus of the protein (data not shown) which could possibly act as a membrane anchor.

The PbpA protein shows homology to a number of low molecular weight penicillin binding proteins (PBP), ranging from PBPs 4 through 7, when analyzed by gBLASTP database searches. The strongest homologs were a protein from *M. tuberculosis* (direct submission) homologous to PBP7 from *E. coli*, a PBP5 from *P. aeruginosa* (direct submission), the DacF protein from *B. subtilis* (direct submission) and a probable PBP from *E. coli* (direct submission). Unexpectedly, the Pbp from the *N. lactamdurans* cluster (51) was not identified as homologous in the database search. This Pbp shows more homology to D-ala-D-ala carboxypeptidases and β-lactamases (data not shown) and shares only 18.8% identical amino acids with the *S. clavuligerus* PbpA.

The PbpA protein may also function as a resistance mechanism to provide protection to the cell from the antibiotic that it produces.

III.5.5 The *cmcT* gene product

The *cmcT* gene encodes a 489 amino acid protein, CmcT, with a calculated molecular weight of 49.7 KD. The hydropathy plot of the CmcT protein showed many transmembrane domains (data not shown) most likely involved in inserting the protein into the cytoplasmic membrane.

The CmcT protein shows homology to a large number of membrane-spanning proteins responsible for export of antibiotics and/or resistance to antibiotics when analyzed by gBLASTP database searches. The strongest homologs were the CmcT protein from *N. lactamdurans* (51), the Pur8 protein responsible for puromycin resistance in *S. anulatus* (210), the EpfA protein, a putative efflux protein from *M. tuberculosis* (direct submission), and the tetracenomycin C resistance and export protein from *S. glaucescens* (88).

The CmcT protein shows homology to the drug resistance translocator family and likely functions as a transporter that exports cephamycin from the cell.

III.5.6 The *pcd* gene product

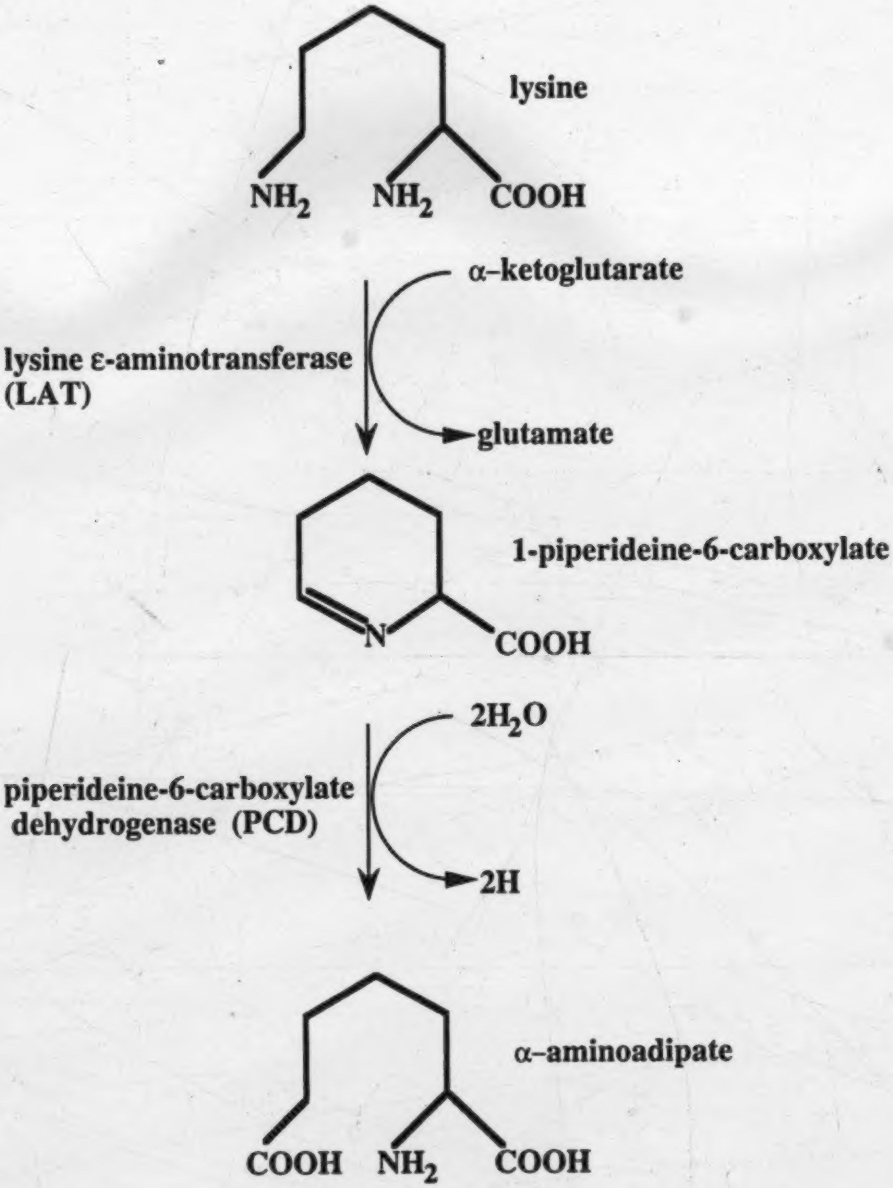
The *pcd* gene encodes a 512 amino acid protein, PCD, with a calculated molecular weight of 54.0 KD. The hydropathy plot of the PCD protein does not show a concentration of hydrophobic amino acids except for the center of the protein (data not shown) suggesting that the protein is cytoplasmic.

The PCD protein shows strong homology to a very large number (>50) of semialdehyde dehydrogenases when analyzed by gBLASTP database searches. The strongest homologs were an aldehyde dehydrogenase homolog from *M. tuberculosis* (direct submission), a hypothetical aldehyde dehydrogenase like-protein from *Caenorhabditis elegans* (direct submission), the turgor-responsive protein from the garden pea, *Pisum sativum*, (87), the antiquitin protein from humans (133), and a turgor-responsive protein from canola, *Brassica napus* (201). The first homologous protein that had been biochemically characterized was the succinate-semialdehyde dehydrogenase from *E. coli* (159). The majority of the homologous proteins are semialdehyde dehydrogenases while the strongest homologs are referred to as aldehyde dehydrogenases even though their activities have not been characterized.

III.5.6.1 Alignment of the PCD protein active site with homologs

The presence of such strong homologs from eukaryotic backgrounds was both surprising and confusing but unfortunately none of these homologous proteins have been biochemically characterized. However, it is likely that PCD functions as an semialdehyde dehydrogenase involved in cephamycin biosynthesis due to its location within the cephamycin cluster. Kern et al. (119) were the first to demonstrate that the production of α-aminoadipate in prokaryotes is a two step process (Fig. 10). The first step is catalyzed by the lysine ε-amino transferase (LAT) protein which removes the epsilon amino group from lysine to create the intermediate 1-piperideine-6-carboxylate (cyclized form of aminoadipic semialdehyde). The second step converting 1-piperideine-6-carboxylate to α-aminoadipate is catalyzed by a piperideine-6-carboxylate dehydrogenase (PCD) which has recently been purified from *S. clavuligerus* (55). Even though a number of gene clusters from bacterial β-lactam producers have been partially or fully sequenced, no candidate genes have been identified (170) and it was thought that the reaction was possibly catalyzed by a primary metabolic enzyme. *pcd* may represent that missing gene, a speculation that is consistent with the location of the *pcd* gene within the cephamycin cluster and with the estimated size of the purified PCD protein at 52.6 KD (55) which corresponds well with the deduced molecular weight of the *pcd* gene product at 54.0 KD.

Figure 10. Pathway for α-aminoadipate biosynthesis in *S. clavuligerus*.


lysine
NH2
NH2
COOH
α–ketoglutarate
lysine ε-aminotransferase
(LAT)
glutamate
1-piperideine-6-carboxylate
N
COOH
2H2O
piperideine-6-carboxylate
dehydrogenase (PCD)
2H
α–aminoadipate
COOH
NH2
COOH

The PCD protein was aligned with some of the most homologous proteins to confirm that it was indeed an semialdehyde dehydrogenase. The deduced amino acid sequences of the PCD homologs described were aligned (Fig. 11) using the ClustalW 1.7 multiple sequence alignment program (212) and the conserved amino acids were shaded using the BOXSHADE 3.21 program. For simplification of the alignment figure only the active sites of the homologous proteins are shown.

Many amino acids were conserved throughout the active site of the proteins, as well many other pockets of highly conserved amino acids existed throughout the proteins (data not shown). A consensus sequence was generated from the alignment of the seven homologous proteins with upper and lower case text designating the identical and functionally similar amino acids, respectively. Aldehyde dehydrogenases have a highly conserved decapeptide, VTLELGGKSP, which is followed by a cysteine 28 amino acids downstream (180). The glutamic acid and cysteine residues are essential for the catalytic activity of the dehydrogenases. The PCD protein has a sequence of TILELGGNNA which has diverged from the aldehyde dehydrogenase consensus sequence but all of the other homologous proteins decapeptides are equally different and all of the proteins also have a cysteine located 28 amino acids downstream. The proteins may differ from the conserved sequence because they are semialdehyde dehydrogenases rather than aldehyde dehydrogenases.

III.5.7 The *cefE* gene product

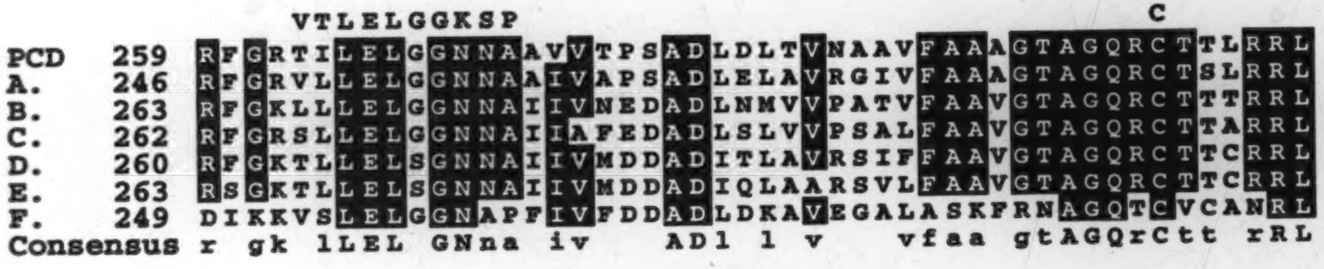
The *cefE* gene encodes a 312 amino acid protein, desacetoxycephalosporin C synthase (DAOCS), with a calculated molecular weight of 34.5 KD. The *cefE* gene was sequenced by Kovacevic et al. (124) and heterologous expression of the *cefE* gene allowed *E. coli* to produce a DAOCS protein that catalyzed the expansion of the thiazolidine ring in penicillin N to a dihydrothiazine ring in desacetoxycephalosporin C .

III.5.8 The *cefD* gene product

The *cefD* gene encodes a 398 amino acid protein, IPNE, with a calculated molecular weight of 43.5 KD. The *cefD* gene was sequenced along with *cefE* by Kovacevic et al. (123) and the IPNE protein was purified from the cytoplasm and found to isomerise isopenicillin N to penicillin N (217).

Figure 11. Alignment of proteins homologous to PCD. A multiple alignment of homologous proteins was done using the ClustalW 1.7 program and displayed using the BOXSHADE 3.21 program. Only the region around the active site of the dehydrogenase homologous proteins was examined. Conserved amino acids were printed in white on a black background and functionally similar amino acids were printed in black within a shaded background. Gaps, marked by dashes, were introduced to reveal the maximal similarity among the sequences. A consensus sequence was given below the alignment and the consensus active site sequence for aldehyde dehydrogenases (180) is given above the alignment.

Protein A. is from *M. tuberculosis*, protein B. is from *Caenorhabditis elegans*, protein C. is from *Pisum sativum*, protein D. is from *Homo sapiens*, protein E. is from *Brassica napus* and protein F. is from *E. coli*.



```
               VTLELGGKSP                           C
PCD    259 RFGRTILELGGNNAAVVTPSADLDLTVNAAVFAAAGTAGQRCTTLRRL
A.     246 RFGRVLLELGGNNAAIVAPSADLELAVRGIVFAAAGTAGQRCTSLRRL
B.     263 RFGKLLLELGGNNAIIVNEDADLNMVVPATVFAAVGTAGQRCTTTRRL
C.     262 RFGRSLLELGGNNAIIAFEDADLSLVVPSALFAAVGTAGQRCTTARRL
D.     260 RFGKTLLELSGNNAIIVMDDADITLAVRSIFFAAVGTAGQRCTTCRRL
E.     263 RSGKTLLELSGNNAIIVMDDADIQLAARSVLFAAVGTAGQRCTTCRRL
F.     249 DIKKVSLELGGNAPFIVFDDADLDKAVEGALASKFRNAGQTCVCANRL
Consensus  r gk lLEL GNna iv   ADl l v   vfaa gtAGQrCtt rRL
```

III.5.9 The *cmcI* gene product

The *cmcI* gene encodes a 236 amino acid protein, CmcI, with a calculated molecular weight of 27.6 KD. The CmcI protein was purified as a complex with CmcJ from the cytoplasm of *N. lactamdurans* (70).

The CmcI protein only shows significant homology to the CmcI protein from *N. lactamdurans* .

The CmcI protein in the CmcI:CmcJ complex of *N. lactamdurans* functions as a cephalosporin-7-α-hydroxylase converting *O*-carbamoyl-desacetylcephalosporin C (OCDAC) to 7-α-hydroxy-OCDAC (70) and CmcI likely functions in a similar manner in *S. clavuligerus*.

III.5.10 The *cmcJ* gene product

The *cmcJ* gene encodes a 310 amino acid protein, CmcJ, with a calculated molecular weight of 34.0 KD. The CmcJ protein was purified as a complex with CmcI from the cytoplasm of *N. lactamdurans* (70).

The CmcJ protein only shows significant homology to the CmcJ protein from *N. lactamdurans* and ORF353 of unknown function from *Methylobacterium* sp. (131).

The CmcJ protein in the CmcI:CmcJ complex of *N. lactamdurans* functions as a methyltransferase converting 7-α-hydroxy-OCDAC to cephamycin C (70) and CmcJ likely functions in a similar manner in *S. clavuligerus*.

III.5.11 The *cefF* gene product

The *cefF* gene encodes a 318 amino acid protein, DACS, with a calculated molecular weight of 34.6 KD. The gene was sequenced by Kovacevic and Miller (122) and the purified desacetylcephalosporin C synthase (DACS) protein was found to catalyze the hydroxylation of desacetoxycephalosporin C to desacetylcephalosporin C (B2).

III.5.12 The *cmcH* gene product

The *cmcH* gene encodes a 522 amino acid protein, CmcH, with a calculated molecular weight of 57.3 KD. The hydropathy plot of the CmcH protein lacks a concentration of hydrophobic amino acids (data not shown) suggesting that the protein is cytoplasmic. The incomplete sequence for the *cmcH* gene from *S. clavuligerus* was described (54) but was not deposited into the databases.

The CmcH protein shows strong homology to nodulation proteins from *Rhizobium* spp. when analyzed by gBLASTP database searches. The strongest homologs were the CmcH protein from *N. lactamdurans* (54) and the NodU proteins from *Rhizobium tropici* (220), *Rhizobium* sp. NGR234 (102) and *Rhizobium fredii* (125). The NodU proteins are *O*-carbamoyltransferases involved in the synthesis of nodulation factors (183).

The CmcH protein from *N. lactamdurans* was believed to be a 3'-hydroxymethylcephem *O*-carbamoyltransferase (54) and likely has the same function in *S. clavuligerus*.

III.5.13 The *ccaR* gene product

The *ccaR* gene encodes a 256 amino acid protein, CcaR, with a calculated molecular weight of 28.6 KD. The hydropathy plot of the CcaR protein lacks a concentration of hydrophobic amino acids (data not shown) suggesting that the protein is cytoplasmic. Further characterization of the CcaR protein is described in section III.6.

III.5.14 The *orf11* gene product

The *orf11* gene encodes a 264 amino acid protein, Orf11, with a calculated molecular weight of 54.0 KD. The hydropathy plot of the Orf11 protein lacks a concentration of hydrophobic amino acids (data not shown) suggesting that the Orf11 protein is cytoplasmic.

The Orf11 protein shows no significant homology to any known proteins when analyzed by gBLASTP or PSI-BLAST database searches.

III.5.15 The *blp* gene product

The *blp* gene encodes a 182 amino acid protein, Blp, with a calculated molecular weight of 20.3 KD. The hydropathy plot of the Blp protein showed a single stretch of hydrophobic amino acids in the N-terminus of the protein (data not shown) which could possibly act as a membrane anchor, similar to the PbpA protein or as a signal sequence like the Bla protein. Perez-Llarena et al. (174) suggested that Blp protein has a precursor form which is processed to remove the 28 amino acid leader peptide. The signal peptide processing site predicted for Blp and BLIP were confirmed by a signal peptide predication program (160).

III.5.15.1 Alignment of Blp protein with BLIP

The Blp protein shows weak homology to only the BLIP protein, also from *S. clavuligerus* (62). The BLIP protein was shown to be an extracellular protein which acts as a β-lactamase inhibitor. The X-ray crystallographic structure of BLIP and the BLIP:TEM β-lactamase complex has been determined (203, 204) and the majority of the

BLIP amino acid residues involved in the protein:protein interactions with the TEM β-lactamase are not the amino acid residues conserved between the BLIP and Blp proteins (Fig. 12).

The function of the Blp protein in *S. clavuligerus* cannot be determined from the database searches or from the results of the BLIP studies.

III.5.16 The *lat* gene product

The *lat* gene encodes a 457 amino acid protein, LAT, with a calculated molecular weight of 54.0 KD. The LAT protein was purified from the cytoplasm and found to act as an aminotransferase catalyzing the conversion of lysine to 1-piperideine-6-carboxylate (119, 188).

III.5.17 The *pcbAB* gene product

The *pcbAB* gene from *S. clavuligerus* has only been partially sequenced but the *pcbAB* gene from *N. lactamdurans* has been completely sequenced and shown to encode a 3649 amino acid protein, δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase (ACVS), with a calculated molecular weight of 403.8 KD (52). The ACVS protein has been purified from the cytoplasm of *S. clavuligerus* and found to function as a peptide synthetase catalyzing the ligation of α-aminoadipate, cysteine and valine into the δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV) tripeptide (117).

III.5.18 The *pcbC* gene product

The *pcbC* gene encodes a 329 amino acid protein, IPNS, with a calculated molecular weight of 36.9 KD. The gene was sequenced by Leskiw et al. (134) and the IPNS protein was purified from the cytoplasm of *S. clavuligerus* and found to catalyze the cyclization of the ACV tripeptide (110).

III.5.19 The *pcbR* gene product

The *pcbR* gene encodes a 551 amino acid protein, PcbR, with a calculated molecular weight of 57.3 KD. The *pcbR* gene was sequenced and characterized by Paradkar et al. (166) and the PcbR protein was shown to be associated with the cell membrane. The hydropathy plot of the PcbR protein showed a single stretch of hydrophobic amino acids in the N-terminus of the protein (data not shown) which could possibly act as a membrane anchor, very similar to the PbpA protein.

The PcbR protein shows weak homology to a number of low affinity PBPs of higher molecular weight (1 through 5) and probably functions as a resistance mechanism

Figure 12. Alignment of BLIP and the Blp protein. The two proteins were aligned using the ClustalW 1.7 program. Conserved amino acids are shown between the two protein sequences while functionally similar amino acids are marked with a +. The putative signal sequences are underlined and the residues important for protein:protein interaction between BLIP and the TEM β-lactamase (203) are marked with dots.

```
Blp           MVKKTWRSALALTAAAGAVLATATSAHAYTGFTPERYNKIQFGMDRTLVWQLAGAD--Q
               V++  R+++  +A++G VL++A +++A   +T  ++++IQFGM R  V ++AGA+  +
BLIP  MRTVGIGAGVRRLGRAVVMAAAVGGLVLGSAGASNAAGVMTGAKFTQIQFGMTRQQVLDIAGAENCE
                                                                      •

Blp   SCSDQVERIICYNNP--DHYGPQGHFFFN-AADKLIHKRQMELFPAPK-PTMRLATYNKTQTGMTEA
      + ++  + I C ++   D+Y+     F + AAD  + ++++E + AP+ PT+ LA +N++ +GMT+A
BLIP  TGGSFGDSIHCRGHAAGDYYAYATFGFTSAAADAKVDSKSQEKLLAPSAPTLTLAKFNQVTVGMTRA
         ••  •   • • ••••                    • ••

Blp   QFWAAVPSDTCSALAEQYPNWPATNGNLREYVC-----PSKAERFAPSAYFTFTDGKLTSRSQSQLP
      Q  A+V + +C++ +E YP +P+T G      C      S++  +  SA++ FTDG L ++ Q +L
BLIP  QVLATVGQGSCTTWSEYYPAYPSTAGVTLSLSCFDVDGYSSTGFYRGSAHLWFTDGVLQGKRQWDLV
                   ••                           ••       •         • •
```

to provide protection to the cell from the antibiotic that it produces (166). PcbR mutants show decreased resistance to benzylpenicillin and cephalosporins.

III.6 Characterization of the CcaR protein

The overall objective of this study was to understand more about how production of cephamycin biosynthesis is regulated in *S. clavuligerus*. The *ccaR* gene is the only obvious regulatory gene located within the cephamycin cluster.

III.6.1 Detection of proteins homologous to CcaR

A number of proteins homologous to CcaR were detected (Table 9) using gBLASTP protein database searches (9, 10). The majority of these proteins were transcriptional activators from *Streptomyces* involved in the regulation of antibiotic biosynthesis. The homology of the DclX/CcaR protein with these transcriptional activators was first mentioned by Barton et al. (20) and by Perez-Llarena et al. (174). Since the completion of these investigations the sequences of more homologous proteins have been added to the databases. Surprisingly, the most similar proteins were not the *Streptomyces* transcriptional activators but rather were proteins of unknown function from *Mycobacterium*.

The transcriptional activators DnrI (206), ActII-ORF4 (74), RedD (155) and SnoA (232) are responsible for regulation of the biosynthesis of daunorubicin, actinorhodin (ACT), undecylprodigiosin (RED) and nogalamycin, respectively. The AfsR protein plays a more global role by regulating both ACT and RED biosynthesis in *S. coelicolor* (77). Of the *Mycobacterium* homologs the EmbR protein appears to be involved in cell wall biosynthesis (22), while Cy50.15 and Y164 are homologous ORFs that came from the *M. tuberculosis* genome sequencing project and have unknown function (direct submissions).

III.6.2 Alignment of CcaR with homologous proteins

The deduced amino acid sequences of the CcaR homologs were aligned (Fig. 13) and the conserved amino acids were shaded. For simplification of the alignment figure only the regions homologous with the CcaR protein were included since many of the proteins were larger than CcaR. The RedD, SnoA and AfsR amino acid sequences began at amino acid 67, 17 and 14 respectively. Most of the homologous proteins also extended beyond the end of the CcaR protein.

Alignment of the homologous proteins demonstrated that many identical and similar amino acids were conserved throughout the entire length of the proteins. Highly

TABLE 9. Proteins showing significant homology to the CcaR protein

Protein	Organism	Regulated biosynthesis	gBLASTP scores[a]	Percent identity[b]	Length[c]
CcaR	*S. clavuligerus*	cephamycin clavulanic acid	-	-	256
EmbR	*M. avium*	cell wall	111, 5e-24	33.2 (256)	384
Y164	*M. tuberculosis*	unknown	104, 4e-22	29.6 (243)	289
DnrI	*S. peucetius*	daunorubicin	103, 1e-21	30.8 (266)	272
Cy50.15	*M. tuberculosis*	unknown	103, 1e-21	31.0 (258)	388
AfsR	*S. coelicolor*	actinorhodin undecylprodigiosin	91.3, 4e-18	29.6 (243)	993
SnoA	*S. nogalater*	nogalamycin	89.5, 2e-17	30.2 (245)	665
ActII-ORF4	*S. coelicolor*	actinorhodin	89.5, 2e-17	30.9 (256)	255
RedD	*S. coelicolor*	undecylprodigiosin	88.3, 3e-17	33.9 (257)	350

[a] The homologous proteins were ranked in order with respect to their gBLASTP scores (MSP, *E*-value) calculated during the database searches.

[b] The percent identity of each homologous protein to the CcaR protein was calculated using the LFASTA program over a length of amino acids, shown in brackets, specific for each protein.

[c] Length of the homologous protein in amino acids.

Figure 13. Alignment of proteins homologous to CcaR. A multiple alignment of the homologous proteins was done using the ClustalW 1.7 program and displayed using the BOXSHADE 3.21 program. Only the region of homology was aligned so the extra amino acids from the larger proteins were removed. Conserved amino acids were printed in white on the black background and functionally similar amino acids were printed in black within the shaded background. A consensus sequence was given below the alignment. The asterisk denotes the TTA encoded leucine of CcaR that is conserved throughout the transcriptional regulators. The two asterisks denote the two adjacent conserved prolines.

```
                                                            *                     **
CcaR          1 -----------VTIRLLGPVTLVKG--SVPIPIRGQRQRRFLASLALRPGQVISKEAIIEDSWDGEPPLTV
EmbR          1 -MANAG-----LGFGVLGPLSVTAG--GAQLPLGAPKQRAVLAMLLIHRNRPVSVDALIDGVWGAAPVPAA
Y164          1 -MAGSATVEKRLDFGLLGPLQMTID--GTPVPSGTPKQRAVLAMLVIMRNRPVGVDALITALWEEWPPSGA
DnrI          1 -----------MQINMLGPLVAHHN--GTSVTPIARKPRQVFSLLALQAGTVVPVPALMDELWGTQPPASA
Cy50.15       1 -----------MQFNVLGPLELNLR--GTKLPLGTPKQRAVLAMLLLSRNQVVAADALVQAIWEKSPPARA
AfsR         14 GFPAEEEESGALRFGVLGPVRAWRD--GETLATGSPQQRALLAALLLREGRTATAGELIDALWGEEPPSQA
SnoA         17 GASSKEKREEDLEFRLLGPVEILWQ--GRNIMPTAPKPRQVISLLMLRHNTVVQAAELIDELWPELPPPSA
ActII-ORF4    1 -----------MRFNLLGRVHVITDA-GCVIPLKSSKATQLLVLLLLRRHEVVGSGVLIEELWADHPPRSA
RedD         67 HGPSLYNFGGCVEINILGPVSIDTSHSGGGIRAG--KVRTLVATLAIDAGRAVSLADLVDELWGATPPDNV
Consensus                  v   lLGpv       g i      k r vl  L l     v    li  vW   Pp

CcaR         59 SGQLQTSAWMIRTALAEAGLPRDA------LGSHDRGYELRVLPD-SIDLFVFREAVRAVRDLHARGQHQ
EmbR         64 RTSIQSYVSTLRRLLRSALPNPNG-----LLASVPPGYQLNVADA-DCDLGRFRGHKTAGVQAASRGRFE
Y164         69 RASIHSYVSNLRKLLGGAGIDPRV-----VLAAAPPGYRLSIPDN-TCDLGRFVAEKTAGVHAAAAGRFE
DnrI         59 LTTLQTYILQVRRGITVALGASHNGPAKDVLRTCYGGYLLDVDPT-NTDVYAFERLAEEGKRACERGELD
Cy50.15      59 RRTVHTYICNLRRTLSDAGVDSRN-----ILVSEPPGYRLLIGDRQQCDLDRFVAAKESGLRASAKGYFS
AfsR         84 LAAVRTYASRLRKVLDPGV-----------LVSESGGYAVRGLAEGALDLARAQDLASAAEKARSAGDLC
SnoA         87 ITTLQTYIYKFRKILMKQG--ADD-----LLRTQPGGYILTIPPS-TVDVNRFERDADDGQELLQRG-DA
ActII-ORF4   60 MTTLQTYVYHTRRLLGEHRVTSDD---RELVLTQPPGYFALIDED-ELDVAVAERLIRTGGRLLEENRLE
RedD        137 LNALQAHAARARKVLNERACPERAG---GILRSVLGGYLLEIDPQ-CVDGNRFLRLVSQGAALLPAD-PT
Consensus          l t     Rr l               l s   GY l i      Dl         g

CcaR        122 EASERLDTALALWKGPAFADVT-SSRLRLRGETLEEERTAAVELRALIDVGLGYYGDAITRLSELVDHDP
EmbR        128 EASGHLAAALREWRGPVLDDLHNFAFVDAFANLLLEERVAAHTARAEAEIACGRADGVIGELEALVAQHP
Y164        133 QASRHLSAALREWRGPVLDDLRDFQFVEPFATALVEDKVLAHTAKAEAEIACGRASAVIAELEALTFEHP
DnrI        128 LASARFRQALDLWRGDALVDVHAGMRIGMEVARLEESRLGVLEARMETDLRLGRHAGLLPELSALTARHP
Cy50.15     124 EAIRYLDSALQNWRGPVLGDLRSFMFVQMFSRALTEDELLVHTKLAEAAIACGRADVVIPKLERLVAMHP
AfsR        143 HARDLLRRALDLWDGEVLAGVP-GPYAQTQRVRLGEWRLQLEETRLDMDLDQGCHAEAVSELTALTAAHP
SnoA        148 AGGTKLGHALALWRGPALADVVASGRLFSYVTRLEEELRFRILELRIEADLATGRHRELVSELKSLVLAHP
ActII-ORF4  126 EALASLDAGLDLWRGPALSTVPCGRVLESNIAHLEELRLFGMQLRIDANWRLGRIGPMIPELRSLVISHP
RedD        203 RAVELLETGLRLWRGPALIDAGEGRRCRGAAALFEERRLTALEDLISAMFLRGGEAQAIAMLQQLVAQYP
Consensus        a   l  aL  WrG  l  v            l E r              G     i  L  L    P

CcaR        191 FREDLYVSLMKAYYAEGRQADAIQVFHRAKDILREQIGISPGERMTRVMQAILRQDEQVLRVGTPA-
EmbR        198 YREPLWAQLITAYYVTERQSDALGAYRRLKTALAEGLGIDPGPTVKALQQRILRQEPLGPGQPGKAA
Y164        203 YREPLWTQLITAYYLSDRQSDALGAYRRVKTTLADDLGIDPGPTLRALNERILRQQPLDAKKSAKTT
DnrI        198 MHENLWAQFMIALHRSGRTSQALEAFIKLRKTLVNELGVEPSARLQHLQHAILRADPGIDRNGPEVP
Cy50.15     194 YRESLWKQLMLGYYVNEYQSAAIDAYHRLKSTLAEELGVEPAPTIRALYHKILRQLPMDDLVGRVTR
AfsR        212 LRERLRELLMLALYRSGRQAEALAVYADTRRLLADELGVDPRPGLRELQQRILQADPALAELSAATA
SnoA        218 LHEHLHGLLMLALHRSGRPHEALEVYRSVRHKMIEDLALEPAQDFATLHHTLLSDSPPEAPEPLWPA
ActII-ORF4  196 LNETLHAKLMGALCQMGRRAEALESYRNLRRILSDELGVDPTPEIQRMHMEILNGEKVLV-------
RedD        272 LRERFCELLMVGLYRVGRQGDALESYRLARKRLDDELGVQPGALLRRRHAEILAQDPVLKVPSALWR
Consensus        rE l   lm a     r   Al  y   r  l e lgv P   l  l   iL
```

conserved sequences were shown with upper and lower case text designating conserved or functionally similar amino acids in eight of the nine proteins, respectively.

The CcaR protein had further similarity to many of the pathway-specific transcriptional activators from *Streptomyces* spp. in that it also contained the rare leucine encoding TTA codon located near the N-terminal end of the protein. The TTA codon has been demonstrated to be involved in the hierarchy of regulation for antibiotic biosynthesis and sporulation in *S. coelicolor* (136). The biosynthesis of both ACT and RED are post-transcriptionally regulated by the *bldA* tRNA due to the presence of TTA codons in the ActII-ORF4 (74) and RedZ (225) genes, respectively. Interestingly, the leucine encoded by the rare TTA codon at position 32 of the CcaR protein was one of the amino acids conserved throughout the CcaR homologs. As well, two adjacent proline residues at positions 54 and 55 were conserved in the homologs and become important in one of the NTG mutants characterized (see V.8.5).

III.6.3 Estimation of DclX/CcaR start codon

The sequences of Walters et al. (*dclX*, 221) and Perez-Llarena et al. (*ccaR*, 174) are identical within the coding region of the protein but they have chosen different start codons (Fig. 14A). Walters et al. (221) have chosen an ATG codon to give a protein of 262 amino acids and a molecular weight of 29.3 KD while Perez-Llarena et al. (174) have chosen the GTG start codon six amino acids further down the sequence to give a protein of 256 amino acids and a molecular weight of 28.6 KD. Neither of the potential Shine-Dalgarno (S/D) sequences are a close match for the 3' end of the 16S rRNA 3' ...UCUUUCCUCCACUAG...5' from *S. lividans* (24). The DclX S/D was located 8 nts from the initiator codon and the CcaR S/D was located 6 nts away. The range varies from 5 to 12 nts with an average of about 8.5 nts (202) so both S/D sequences fit within the range.

Both the DclX and CcaR versions of the protein were included in an alignment of the N-terminal ends of the proteins for a direct comparison to attempt to determine the most likely start codon for the CcaR protein based upon homology (Fig. 14B). The N-termini of the DnrI, ActII-ORF4 and Y164 proteins all align directly with the CcaR GTG while the EmbR and Cy50.15 proteins begin around the DclX ATG. It seems that the CcaR GTG is the more likely of the two start codons based upon homology although N-terminal sequencing of the purified protein would be required to give a definitive answer.

III.6.4 Detection of structural motifs in the CcaR sequence

The deduced amino acid sequence of the CcaR protein was searched for structural

Figure 14. Estimation of the correct start codon for CcaR.

A. The DclX start codon predicted by Walters et al. (221) and the CcaR start codon predicted by Perez-Llarena et al. (174) were compared. The predicted S/D sequences are given in bold.

B. Alignment of the first 30 amino acids was shown to observe the start codon positions of the CcaR homologs.

A.

DclX

```
ttcacaaggggaccgccATG AAC ACC TGG AAT GAT GTG ACG ATC
                  M   N   T   W   N   D   V   T   I
```

CcaR

```
ttcacaaggggggaccgcc atg aac acc tgg aat gat GTG ACG ATC
                                            V   T   I
```

B.

```
DclX           1  -----MNTWNDVTIRLLGPVTLVKG-SVPI
CcaR           1  -----------MTIRLLGPVTLVKG-SVPI
EmbR           1  ------MANAGLGFGVLGPLSVTAG-GAQL
Cy50.15        1  -----------MQFNVLGPLELNLR-GTKL
DnrI           1  -----------MQINMLGPLVAHHN-GTSV
Y164           1  -MAGSATVEKRLDFGLLGPLQMTID-GTPV
AfsR          14  GFPAEEEESGALRFGVLGPVRAWRD-GETL
SnoA          17  GASSKEKREEDLEFRLLGPVEILWQ-GRNI
ActII-ORF4     1  -----------MRFNLLGRVHVITDAGCVI
RedD          64  HGPSLYNFGGCVEINILGPVSIDTS-HSGG
Consensus                    m   lLGpv         i
```

motifs, like a helix-turn-helix (HTH) domain, using the MOTIF program in the UWGCG package. The program was unable to detect a HTH domain within the CcaR protein but was able to find a *lysR* homologous HTH motif within the ClaR protein, a putative transcriptional activator located within the clavulanic acid cluster shown to be involved in the regulation of the late clavulanic acid biosynthetic genes (165). The majority of these CcaR homologous transcriptional activators do not contain a recognizable HTH motif that would represent the protein structure that mediates site-specific binding of proteins to DNA (150). However, both the RedD and AfsR proteins do possess HTH motifs (99, 155) but the regions predicted to contain these motifs are outside of the CcaR homologous region. Even though the DnrI protein lacks a typical HTH motif it has been demonstrated to possess specific DNA-binding capabilities (209). Recently, Wietzorrek and Bibb (228) suggested that these *Streptomyces* activators along with the EmbR protein constitute a novel family of proteins that contain an OmpR-like DNA-binding fold. They propose that this fold, located towards the N-terminus, resembles the C-terminally located DNA-binding domain of OmpR in both primary amino acid sequence and predicted secondary structure and would allow these proteins to bind heptameric repeats within the promoter proximal region of regulated genes.

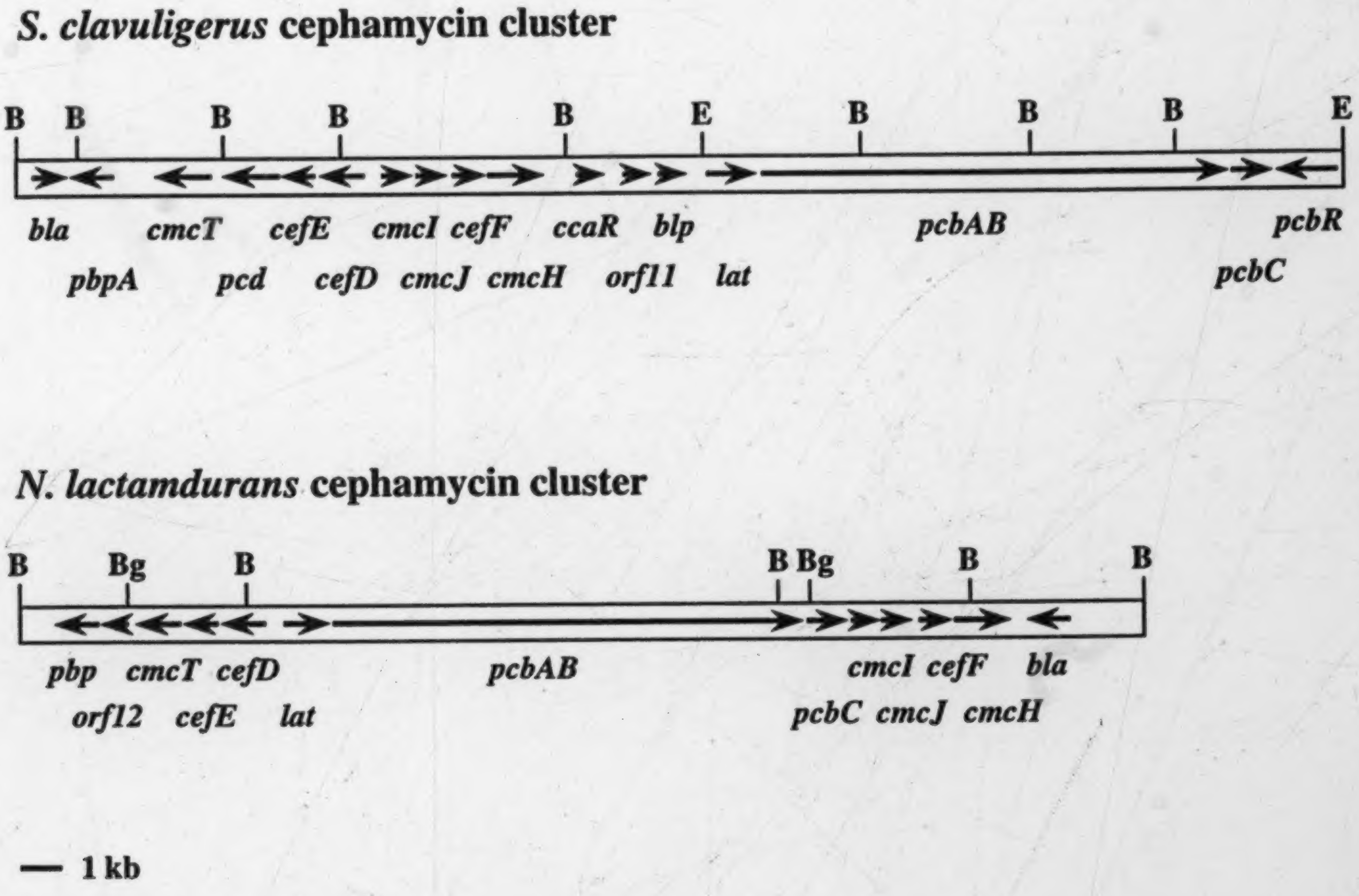
III.7 Comparison of the *N. lactamdurans* and *S. clavuligerus* cephamycin biosynthetic clusters

The cephamycin cluster of *S. clavuligerus* generated from the sequence assembled in this study together with the sequences of *pcbC* (63, 134) and *pcbR* (166) genes was compared to the cephamycin cluster of *N. lactamdurans* (Fig. 15). At first glance the clusters seem quite similar as the majority of the genes present in the *S. clavuligerus* cluster were also present in the smaller *N. lactamdurans* cluster. The most obvious difference was the complete absence of the *ccaR*, *orf11* and *blp* group of genes from the *N. lactamdurans*. Also absent from the *N. lactamdurans* cluster were the *pcbR* and *pcd* genes. Only one gene was present in *N. lactamdurans* that was absent from the *S. clavuligerus* and that was the *orf12* gene. The *orf12* gene was described as a small gene of unknown function (51) but the *orf12* gene sequence has also been deposited into Genbank where it was described as a possible DNA binding protein. The *orf12* gene shows homology to only itself in gBLASTP database searches (data not shown).

Apart from differences in the genes which are present/absent in each cluster, the clusters also differ in the organization of the genes within. Both clusters are flanked by resistance genes. The *S. clavuligerus* cluster has two presumptive resistance genes on one end (*bla* and *pbpA*) and the unique *pcbR* on the other end, while the *N. lactamdurans*

Figure 15. Comparison of the cephamycin biosynthetic clusters from *S. clavuligerus* and *N. lactamdurans*. The *N. lactamdurans* cephamycin cluster was described in Coque et al. (43, 51) and this figure attempts to reproduce the proportions of their restriction map.

Abbreviations for restriction sites: B for *Bam*HI, Bg for *Bgl*II and E for *Eco*RI.


S. clavuligerus cephamycin cluster
B
B
B
B
B
E
B
B
B
E
bla
pbpA
cmcT
pcd
cefE
cefD
cmcI
cmcJ
cefF
cmcH
ccaR
orf11
blp
lat
pcbAB
pcbC
pcbR
N. lactamdurans cephamycin cluster
B
Bg
B
B
Bg
B
B
pbp
orf12
cmcT
cefE
cefD
lat
pcbAB
pcbC
cmcI
cmcJ
cefF
cmcH
bla
1 kb

cluster has the *pbp* gene on one end and the *bla* gene on the other end. The entire group of late genes (*cmcI*, *cmcJ*, *cefF* and *cmcH*) has apparently translocated as a cassette with all genes maintaining the same position and orientation. In *S. clavuligerus* the 'late gene cassette' is located between *cefD* and *ccaR* while in *N. lactamdurans* the 'late gene cassette' is located near the end of the cluster between *pcbC* and *bla*. This places the early genes (*lat*, *pcbAB* and *pcbC*) directly adjacent to the middle genes (*cefD*, *cefE* and *cmcT*) in the *N. lactamdurans* cluster, while in *S. clavuligerus* the early genes are separated from the middle genes (*cefD*, *cefE*, *pcd* and *cmcT*) by the 'late gene cassette' and *ccaR*, *orf11* and *blp*.

III.8 Heterologous probing of actinomycetes

Genomic DNA preparations from a number of actinomycetes were digested to completion with *Nco*I, separated on an agarose gel and subjected to Southern analysis. Gels used to prepare the Southern blots were loaded with equal amounts of DNA as judged by ethidium bromide staining of the gel before transfer. The nylon membranes were hybridized under moderate stringency conditions to look for regions of DNA within these organisms which could cross-hybridize with various fragments of *S. clavuligerus* DNA from the cephamycin cluster. The moderate stringency hybridization and wash conditions used would allow the probe to hybridize to DNA that contains similar but not necessarily identical sequences (see II.2.3.3.1). The hybridization and wash conditions would still be stringent enough to prevent non-specific probe binding. After approximately 24 hours of exposure to X-ray film the first autoradiograph was developed and typically only the *S. clavuligerus* band was seen. The Southern blot was then reexposed to X-ray film for up to 7 days, depending upon the initial strength of hybridization, and this was when the weaker hybridizing bands appeared.

The β-lactam producing actinomycetes used were: *S. clavuligerus* NRRL 3585, *S. cattleya* NRRL 3841, *S. griseus* NRRL 3851, *S. jumonjinensis* NRRL 5741, *S. lipmanii* NRRL 3584 and *Nocardia lactamdurans* NRRL 3802. *S. clavuligerus*, *S. cattleya*, *S. jumonjinensis* and *N. lactamdurans* all produce cephamycin C whereas *S. griseus* produces cephamycins A and B and *S. lipmanii* produces 7-methoxycephalosporin C. These latter three products are non-carbamoylated cephamycins (54). These organisms all produce penicillin/cephalosporin compounds but some also produce additional β-lactam products. *S. clavuligerus*, and *S. jumonjinensis* can also produce clavulanic acid (69) while *S. lipmanii* makes clavaminic acid and precursors (19). In addition, *S. clavuligerus* includes clavams to the number of β-lactam metabolites it makes. A variety of non-β-lactam producing actinomycetes were included: *S. antibioticus* ATCC 8663, *S. fradiae* ATCC 19609, *S. griseofuscus* NRRL B-5429, *S. lividans* 1326 and *S. venezuelae* 13s.

III.8.1 Hybridization with *S. clavuligerus* DNA fragments

The extent of hybridization was given a qualitative score relative to the extent of probe hybridization to the *Nco*I digested *S. clavuligerus* DNA, which obviously would be complementary and was given a score of +++. Moderate hybridization (++) was judged as solid bands which appear after the long exposure but were not seen on the first exposure and weak/very weak hybridization (+) was judged as bands that are still faint after the final exposure. Results of the cross hybridization of other actinomycetes with *S. clavuligerus* DNA probes is summarized in Table 10.

The *lat* probe was an *Nco*I/*Bam*HI fragment containing the *lat* coding region as well as a small amount of the *pcbAB* coding region (nts. 22092-23633) and was used as a positive control for the cephamycin producers (142). The *lat* probe showed hybridization with all β-lactam producing species including *N. lactamdurans* and was absent from the lanes of all non-producers. Equal levels of moderate hybridization were seen in all the other β-lactam producing strains except for *S. jumonjinensis* which had very weak hybridization.

The *ccaR* probe was the *Eco*RI/*Bam*HI PCR fragment used for high level expression in *E. coli* and thus contained the entire *ccaR* gene (*dclX* start codon, nts. 18467-19248) with very little flanking DNA. The *ccaR* probe hybridized to all of the β-lactam producers except for *N. lactamdurans* and did not hybridize to any of the non-β-lactam producers (Table 10 and Fig. 16). Obviously, the *ccaR* probe recognizes itself within the *S. clavuligerus* DNA with the greatest strength but after long exposures discrete bands of hybridization (much weaker) were detected in the lanes of the other *Streptomyces* β-lactam producers.

The *orf11* probe consisted of two consecutive *Sma*I fragments (nts. 19707-20514) which contain almost the entire *orf11* coding region plus a small amount of upstream sequence. The *orf11* probe hybridized only to *S. clavuligerus* (strong hybridization), *S. jumonjinensis* (very weak hybridization) and *S. lipmanii* (moderate hybridization) from the β-lactam producers and hybridized to none of the non-producers.

The *blp* probe was a 680 bp *Not*I/*Pvu*I fragment (nts. 20979-21657) containing most of the *blp* coding region as well as some of the downstream sequences. The *blp* probe hybridized only to *S. clavuligerus* DNA and not to any of the other actinomycetes.

The *pcd* probe was a 1560 bp *Kpn*I/*Bam*HI fragment (nts. 8579-10139) containing almost the entire *pcd* coding region and a small amount of downstream

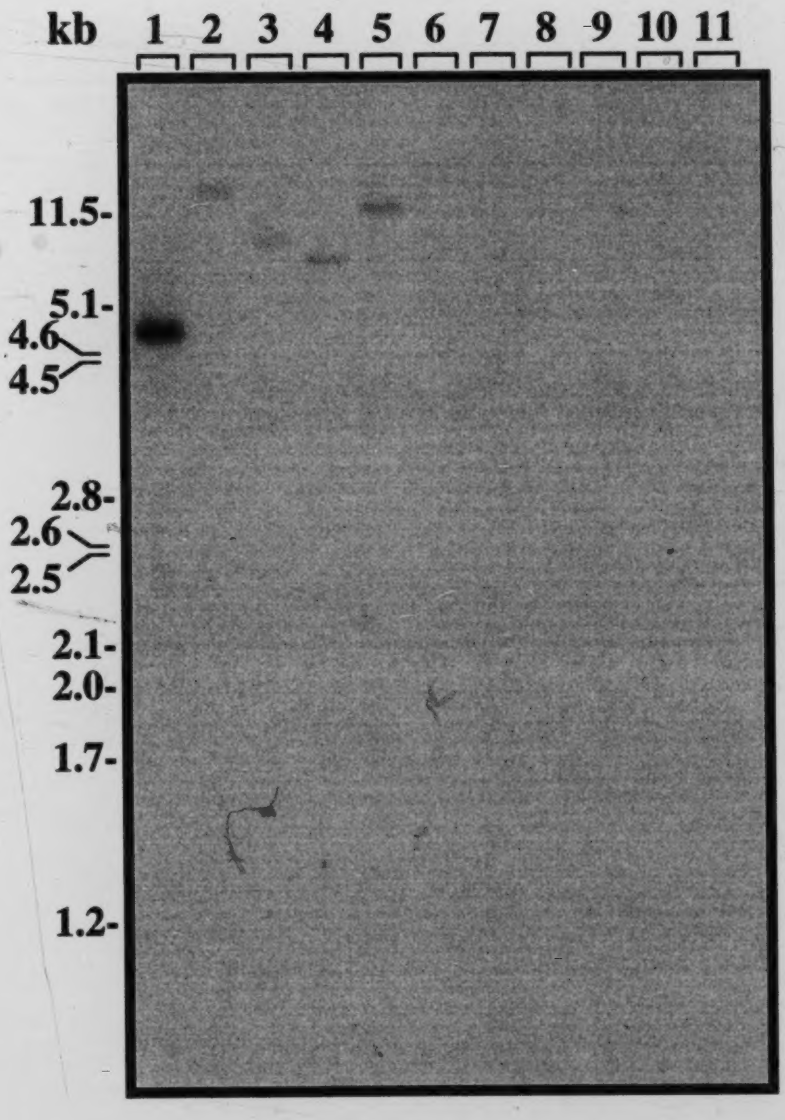
Table 10. Cross hybridization of *S. clavuligerus* cluster genes with other actinomycetes

		S. clavuligerus DNA probes[b]				
Actinomycete[a]	β-lactam antibiotics	*ccaR*	*orf11*	*blp*	*pcd*	*lat*
S. clavuligerus	cephamycin C, clavams, clavulanic acid	+++	+++	+++	+++	+++
S. cattleya	cephamycin C	+	-	-	++	++
S. griseus	cephamycin A and B	+	-	-	++	++
S. jumonjinensis	cephamycin C, clavulanic acid	+	+	-	+	+
S. lipmanii	7-methoxycephalosporin C, clavaminic acid precursors	+	++	-	+++	++
N. lactamdurans	cephamycin C	-	-	-	-	++
S. antibioticus	none	-	-	-	-	-
S. fradiae	none	-	-	-	-	-
S. griseofuscus	none	-	-	-	-	-
S. lividans	none	-	-	-	-	-
S. venezuelae	none	-	-	-	-	-

[a] A combination of β-lactam producing actinomycetes (*S. clavuligerus* NRRL 3585, *S. cattleya* NRRL 3841, *S. griseus* NRRL 3851, *S. jumonjinensis* NRRL 5741, *S. lipmanii* NRRL 3584 and *Nocardia lactamdurans* NRRL 3802) and non-β-lactam producing actinomycetes (*S. antibioticus* ATCC 8663, *S. fradiae* ATCC 19609, *S. griseofuscus* NRRL B-5429, *S. lividans* 1326 and *S. venezuelae* 13s) were used.

[b] The extent of hybridization was a qualitative score relative to the hybridization of the probe to the *S. clavuligerus* genomic DNA ranging from strong hybridization (+++), moderate hybridization (++), weak/very weak hybridization (+) and no hybridization (-).

Figure 16. Cross hybridization of the *ccaR* gene with other actinomycetes. A combination of β-lactam producing actinomycetes (Lanes 1. *S. clavuligerus* NRRL 3585, 2. *S. cattleya* NRRL 3841, 3. *S. griseus* NRRL 3851, 4. *S. jumonjinensis* NRRL 5741, 5. *S. lipmanii* NRRL 3584 and 6. *Nocardia lactamdurans* NRRL 3802) and non-β-lactam producing actinomycetes (lanes 7. *S. antibioticus* ATCC 8663, 8. *S. fradiae* ATCC 19609, 9. *S. griseofuscus* NRRL B-5429, 10. *S. lividans* 1326 and 11. *S. venezuelae* 13s) were used.


kb
1
2
3
4
5
6
7
8
9
10
11
11.5-
5.1-
4.6
4.5
2.8-
2.6
2.5
2.1-
2.0-
1.7-
1.2-

sequence. The *pcd* probe showed strong hybridization to both *S. clavuligerus* and *S. lipmanii* and moderate hybridization to *S. cattleya* and *S. griseus*. *S. jumonjinensis* gave very weak hybridization while *N. lactamdurans* and the non-β-lactam producers failed to hybridize.

IV. Results. Regulation of the *lat*, *pcbAB* and *pcbC* genes

The *lat* gene is present exclusively in prokaryotic β-lactam producers and the lysine ε-aminotransferase (LAT) protein, responsible for the first of a two step reaction converting lysine to α-aminoadipate (αAA), represents the first step of penicillin/cephalosporin biosynthesis (142). Petrich et al. (176) demonstrated that the *lat* promoter appeared to be responsible for the transcription of the *lat*, *pcbAB* and *pcbC* genes, creating a tricistronic message of approximately 14 kb. As well, the *pcbC* gene was shown to have its own promoter located within the *pcbAB* coding region creating a 1.2 kb *pcbC* monocistronic message (178).

Subsequent papers (56, 181, 234) investigating regulation of the *lat* gene gave results which were contradictory to the transcription analysis experiments of Petrich et al. (176, 178) and as a result quite a different model explaining the regulation of the *lat*, *pcbAB* and *pcbC* genes was generated.

A putative promoter within the *lat* coding region gave very strong levels of apparent promoter activity as judged by promoter probe analysis, which led Yu et al. (234) to suggest that this promoter region was responsible for *pcbAB* transcription. It was also suggested that a potential stem-loop structure located in the 152 bp *latpcbAB* intergenic region functioned as a transcription terminator and this structure was preceded by a putative anti-terminator region proposed to allow conditional regulation at the *lat* stem-loop structure (234).

The NP1 mutant discovered and characterized by Piret and coworkers (56, 181, 234) produced very low levels of cephamycin C and LAT activity as well as decreased levels of δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase (ACVS) and isopenicillin N synthase (IPNS) activity. When the NP1 mutant was supplied with αAA exogenously the levels of cephamycin increased suggesting that NP1 was a *lat* mutant. Complementation of the NP1 mutant resulted in increased cephamycin production and LAT activity but unexpectedly the levels of ACVS and IPNS activity also increased.

From these experiments it was concluded that an "interdependence" existed between the first three genes of the penicillin/cephamycin biosynthetic pathway for their transcription (56). In response to these results, Jensen and coworkers (111) created a *lat*::*apr* mutant strain by homologous recombination which was otherwise isogenic to wildtype *S. clavuligerus*. This mutant *lat*::*apr* did not produce any cephamycin but was capable of producing reduced levels of ACVS and IPNS activity and when fed with exogenous αAA produced low amounts of cephamycin. S1 nuclease mapping was unable to determine whether transcription was originating at the *lat* promoter or a promoter located within the *apr* marker and therefore the possibility that the putative promoter located within the *lat* coding region was responsible for *pcbAB* transcription still remained.

Obviously there was considerable interest in understanding the transcriptional organization of the first three genes of the *S. clavuligerus* cephamycin biosynthetic pathway. Characterization of the transcriptional organization and regulation of the *lat*, *pcbAB* and *pcbC* genes was accomplished by promoter probe analysis and gene disruption/replacement methodologies.

IV.1 Construction of promoter probe shuttle vectors

The plasmid pIJ4083 is a high copy number promoter probe vector used for the characterization of promoter fragments in *Streptomyces* (47). The *xylE* reporter gene produced very high levels of catechol 2,3 dioxygenase (C23O) activity in *S. clavuligerus* when the *lat* promoter was inserted (176). The interpretation of results from promoter probe plasmid studies can be complex because the pIJ4083 plasmid itself gives rise to moderate basal levels of C23O activity even though the multiple cloning site (MCS) is preceded by a transcription terminator. Furthermore, insertion of a promoter fragment in the opposite orientation can generate increased levels of C23O activity (164). It was also concerning that the *lat* fragment containing the putative *pcbAB* promoter which provided strong promoter activity in promoter probe experiments did not contain a detectable transcription start point (*tsp*). It suggested that some low level of transcription was occurring in this region which, under normal physiological conditions, would not lead to detectable amounts of transcript but which became active because of the high copy number of the plasmid.

Before the *lat* region was studied by promoter probe analysis changes were made to the existing system relying on the high copy number vector pIJ4083. An effort was made to reduce the background activity from the vector itself by decreasing plasmid copy number and thereby eliminating transcriptional activity not normally detected under physiological conditions. The construction of recombinant pIJ4083-based plasmids begins with transformation of the ligation mixtures into *S. lividans* protoplasts, screening of the transformants for positive promoter-containing constructs and then transforming these constructs into *S. clavuligerus*. By using *E. coli/Streptomyces* promoter probe shuttle vectors it would be possible to generate the constructs rapidly in *E. coli* before moving the final construct into *S. clavuligerus* for analysis. The *E. coli/Streptomyces* promoter probe shuttle vectors have reduced plasmid copy number within *S. clavuligerus* and this hopefully will relieve some or all of these concerns.

IV.1.1 Reconstructing the reporter gene cassette from pIJ4083

The reporter gene cassette, consisting of an fd phage terminator, MCS, *xylE* reporter gene and the lambda t_o terminator, was removed from pIJ4083 as two separate pieces and was reassembled within the vector pIJ2925 (Fig. 17). Firstly, the MCS, *xylE* reporter gene and the t_o terminator were removed as a 1.4 kb *Bgl*II fragment that was inserted into pIJ2925 digested with *Bgl*II to give pDA300. The 1.0 kb *Pst*I fragment containing fd terminator was removed from pIJ4083 and inserted into pBluescript KS(+) digested with *Pst*I to give pDA302. The fd terminator was removed as a *Bgl*II and *Pst*I fragment from pDA302 and inserted into pUC8 digested with *Bam*HI and *Pst*I to give pDA303. The pDA303 plasmid was digested with *Pst*I and *Sph*I to remove a 0.5 kb fragment, the ends were made blunt with the Klenow polymerase and the reaction was self-ligated to give pDA304. The 0.5 kb *Eco*RI and *Hind*III fragment was removed from pDA304 and inserted into pDA300 digested with *Eco*RI and *Hind*III to give pDA305. The insert within the pDA305 vector was the reconstructed reporter gene cassette from pIJ4083 containing only minor alterations to the MCS.

IV.1.2 Construction of a mid-copy number promoter probe vector

A mid-copy number *xylE* promoter probe vector was constructed using the mid-copy number *E. coli/Streptomyces* shuttle plasmid pFD666 (59). The pFD666 plasmid has a pJV1-based origin of replication (ori) for replication in *Streptomyces* and a standard ColE1-based ori for replication in *E. coli.* The vector contains a bifunctional neomycin resistance (*neo*) marker which confers resistance to neomycin or kanamycin for both *E. coli* and *Streptomyces* (58). The mid-copy number vector pDA2500 was constructed from the reassembled reporter gene cassette and the pFD666 vector after minor modifications (Fig. 18).

It was necessary to remove the *Xba*I site from the pFD666 MCS so that the *Xba*I site within reporter gene cassette MCS would be unique. Therefore pFD666 was digested with *Xba*I, the ends were made blunt with Klenow polymerase and then self-ligated to give pDA307. This plasmid was then digested with *Bam*HI and *Hind*III and ligated with the reporter gene cassette removed from pDA305 by digestion with *Bgl*II and *Hind*III to give the vector pDA2500.

IV.1.3 Construction of a single copy promoter probe vector

The single copy *xylE* promoter probe vector was constructed by inserting the *att* fragment which contains the ØC31 attachment site (*attP*) and integrase from pSET152 (27) into a derivative of pFD666. This vector would also contain the *neo* marker for

Figure 17. Reconstruction of the *xylE* reporter gene cassette from pIJ4083 in *E. coli.* The MCS, *xylE* reporter gene and the lambda t_0 terminator were removed as one fragment from pIJ4083 and the fd terminator was removed as another fragment. The two fragments comprising the reporter gene cassette were reassembled in pIJ2925.

Black boxes represent the terminators and the hatched boxes represent the *xylE* reporter gene. The / indicates restriction sites that were destroyed during the ligation of compatible ends that do not regenerate either restriction site and the [] represents sites that were exposed to Klenow polymerase to generate blunt-ended fragments necessary for ligation. Abbreviations for restriction sites: B for *Bam*HI, Bg for *Bgl*II, E for *Eco*RI, H for *Hin*dIII, P for *Pst*I, S for *Sac*I, Sp for *Sph*I and X for *Xba*I.

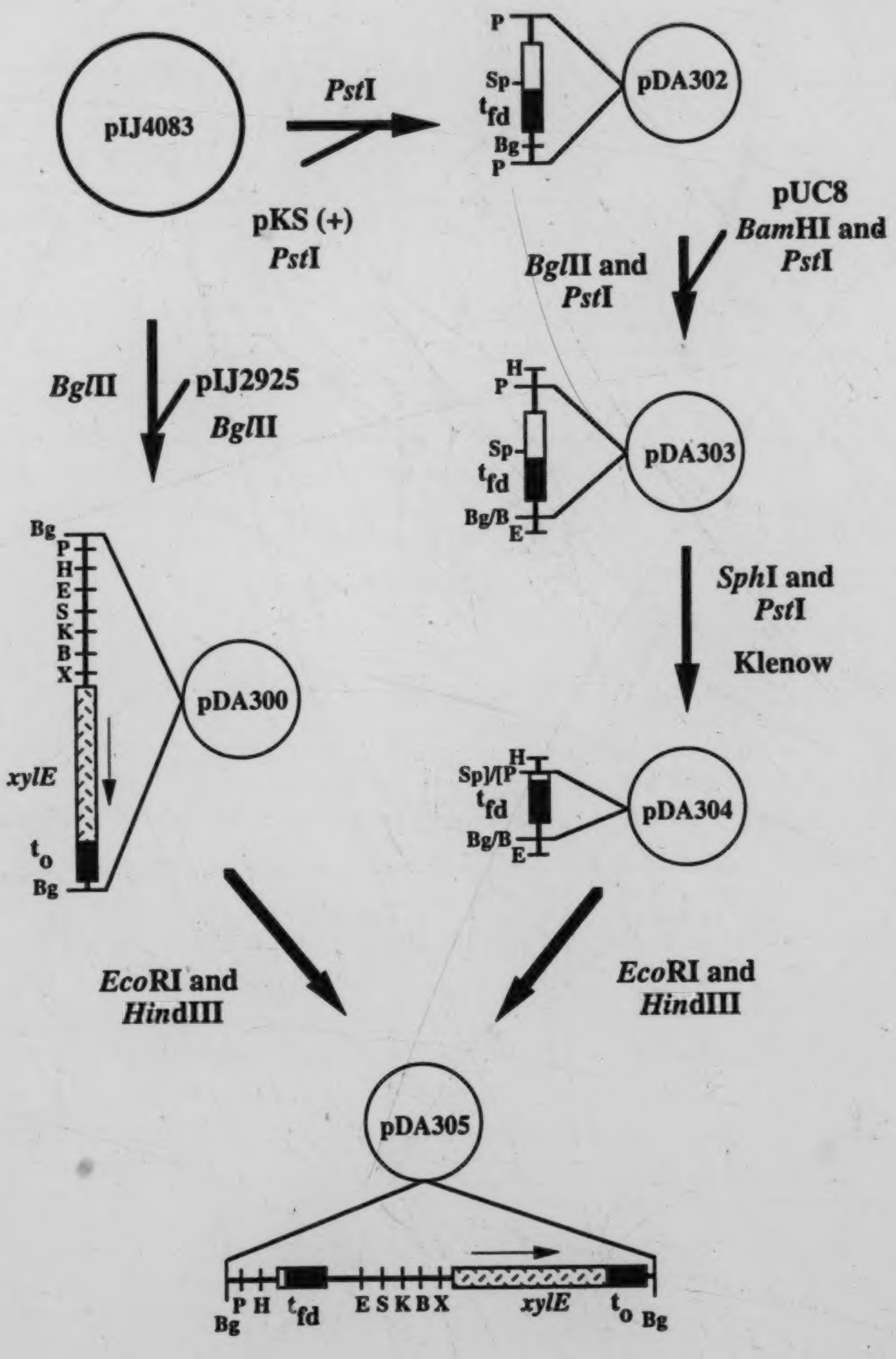
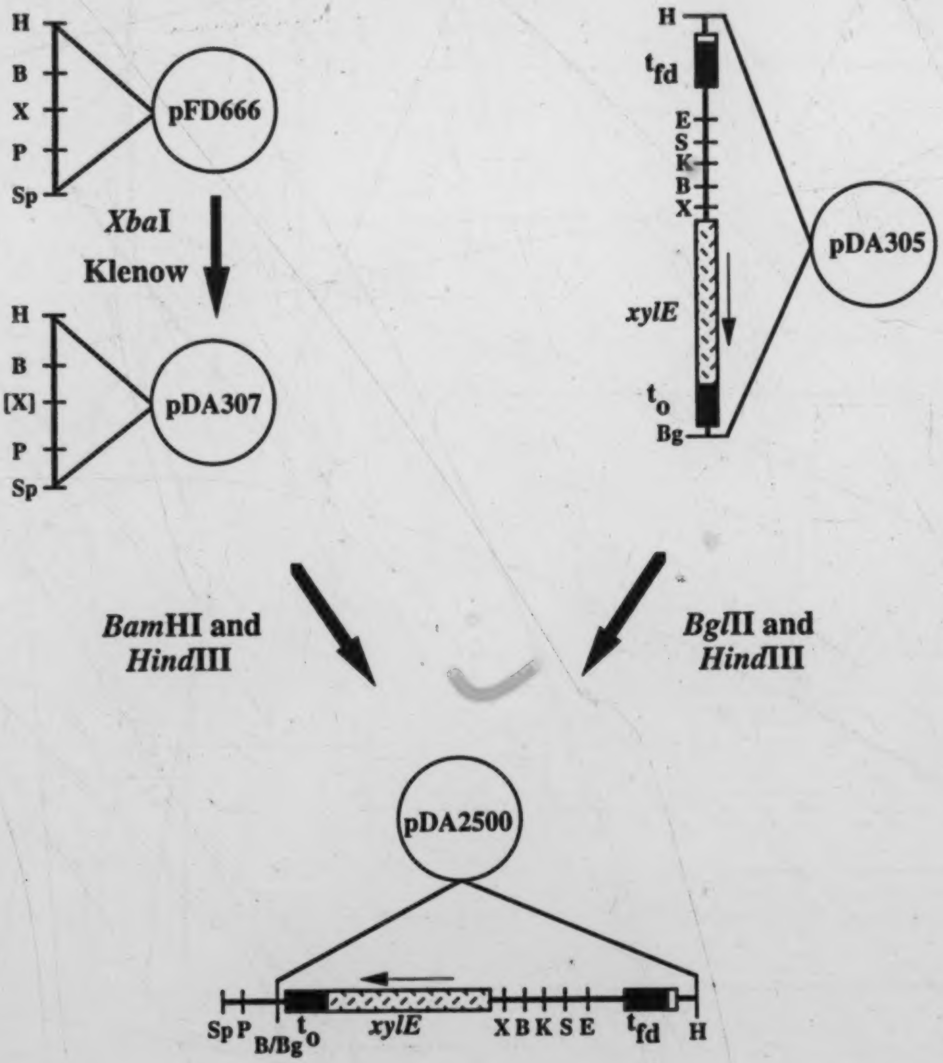

pIJ4083
PstI
pKS (+)
PstI
P
Sp
t_{fd}
Bg
P
pDA302
pUC8
BamHI and
PstI
BglII and
PstI
BglII
pIJ2925
BglII
H
P
Sp
t_{fd}
Bg/B
E
pDA303
SphI and
PstI
Klenow
Bg
P
H
E
S
K
B
X
xylE
t_o
Bg
pDA300
H
Sp]/[P
t_{fd}
Bg/B
E
pDA304
EcoRI and
HindIII
EcoRI and
HindIII
pDA305
Bg
P
H
t_{fd}
E
S
K
B
X
xylE
t_o
Bg

Figure 18. Construction of the mid-copy number promoter probe vector pDA2500. The reconstructed reporter gene cassette was inserted into a modified version of the *E. coli* /*Streptomyces* shuttle plasmid pFD666.

Black boxes represent the terminators and the hatched boxes represent the *xylE* reporter gene. The / indicates restriction sites that were destroyed during the ligation of compatible ends that do not regenerate either restriction site and the [] represents sites that were exposed to Klenow polymerase to generate blunt-ended fragments necessary for ligation. Abbreviations for restriction sites: B for *Bam*HI, Bg for *Bgl*II, E for *Eco*RI, H for *Hin*dIII, P for *Pst*I, S for *Sac*I, Sp for *Sph*I and X for *Xba*I.


H
B
X
P
Sp
pFD666
XbaI
Klenow
H
B
[X]
P
Sp
pDA307
H
t fd
E
S
K
B
X
xylE
t o
Bg
pDA305
BamHI and HindIII
BglII and HindIII
pDA2500
Sp
P
B/Bg
t o
xylE
X
B
K
S
E
t fd
H

selection in both *E. coli* and *Streptomyces*. The single copy vector pDA2000 was constructed from the reassembled reporter gene cassette, the *att* fragment removed from pSET152 and the pFD666 vector after major modifications (Fig. 19).

Again it was necessary to remove the *Xba*I site from the pFD666 MCS so the single copy vector was constructed in pDA307. The pJV1 ori and M13 ori were removed from pDA307 by self-ligation after *Sal*I digestion to remove a 2.7 kb fragment and give the plasmid pDA308. The plasmid pDA308 was linearized with *Sal*I and the adapter oligonucleotide Dyl11 which creates a *Bgl*II site was inserted to create pDA309. This plasmid was then digested with *Bam*HI and *Hind*III and ligated with the reporter gene cassette removed from pDA305 by digestion with *Bgl*II and *Hind*III to give pDA313.

The *att* site was removed as a *Pst*I fragment from pSET152 and inserted into pIJ2925 digested with *Pst*I to create pDA311. To remove the *Eco*RI, *Kpn*I, *Bam*HI and *Xba*I sites from pDA311 the plasmid was digested with *Eco*RI/*Xba*I, the ends were made blunt with the Klenow polymerase and the reaction was self-ligated. The plasmid pDA312 was digested with *Bgl*II and the *att* fragment was inserted into pDA313 digested with *Bgl*II to create pDA2000.

IV.2 Cloning promoter fragments into pDA2000 or pDA2500

Cloning of promoter containing inserts into the pDA2000 and pDA2500 vectors was done in *E. coli* so that potential positive constructs could be rapidly screened. The pDA2500 promoter probe vector would have unique *Eco*RI, *Kpn*I, *Bam*HI and *Xba*I sites within the MCS for cloning the promoter fragments (Fig. 20). The pDA2000 promoter probe vector would have unique *Eco*RI, *Sac*I, *Kpn*I, *Bam*HI and *Xba*I sites within the MCS for cloning the promoter fragments.

The vectors themselves did not display any instability in recombination-deficient or proficient strains of *E. coli* nor did they display any C23O activity when *E. coli* colonies were sprayed with liquid catechol. Most promoter fragments could be inserted into the MCS of either plasmid without difficulty, except for the *ermE** promoter or the *apr* resistance gene, both of which are quite active in *E. coli* and presumably produced toxic amounts of C23O because the plasmids were severely rearranged. It was possible to insert the *ermE** promoter into pDA2500 if the ligation mixture was transformed directly into *S. lividans*. Attempts to transform this construct back into *E. coli* were unsuccessful, suggesting that the instability was due to toxic production of C23O in *E. coli* and not an instability associated with the plasmid itself.

Figure 19. Construction of the single copy promoter probe vector pDA2000. The reconstructed reporter gene cassette and the *att* fragment were inserted into a modified version of the *E. coli /Streptomyces* shuttle plasmid pFD666.

Black boxes represent the terminators and the hatched boxes represent the *xylE* reporter gene. The / indicates restriction sites that were destroyed during the ligation of compatible ends that do not regenerate either restriction site and the [] represents sites that were exposed to Klenow polymerase to generate blunt-ended fragments necessary for ligation. Abbreviations for restriction sites: B for *Bam*HI, Bg for *Bgl*II, E for *Eco*RI, H for *Hin*dIII, P for *Pst*I, S for *Sac*I, Sp for *Sph*I and X for *Xba*I.

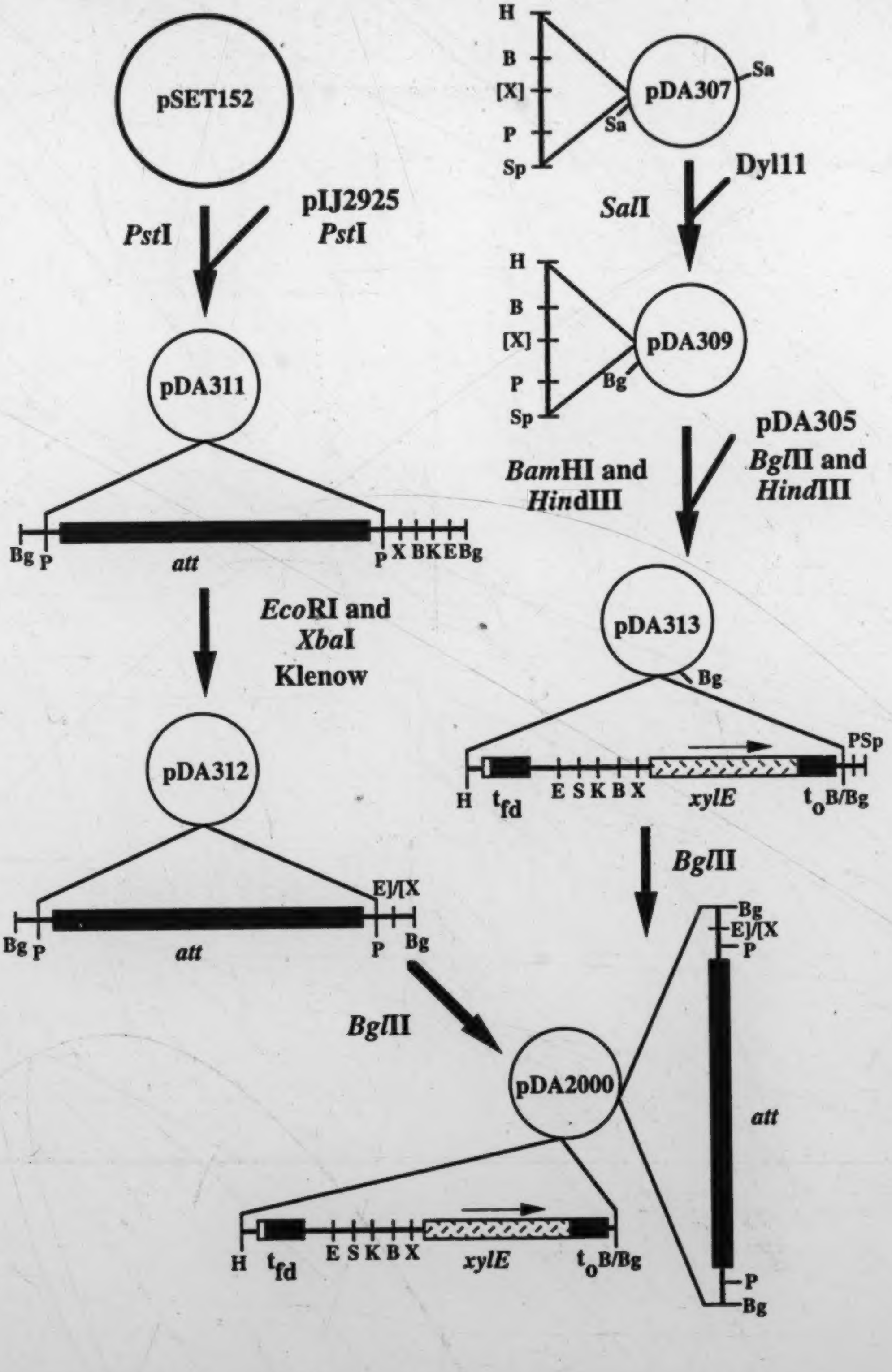
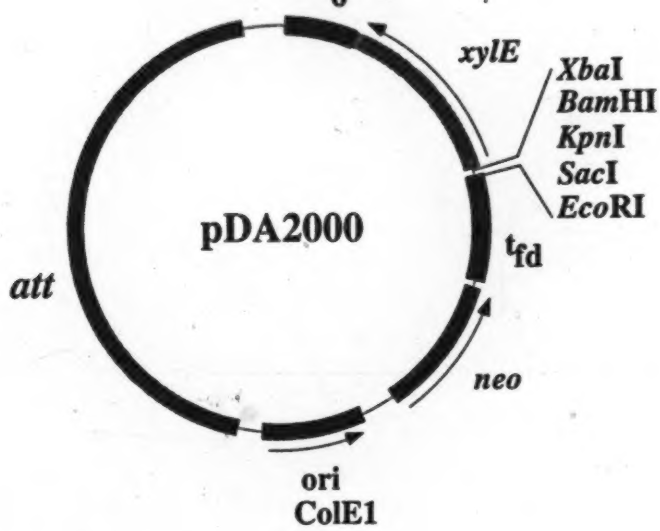
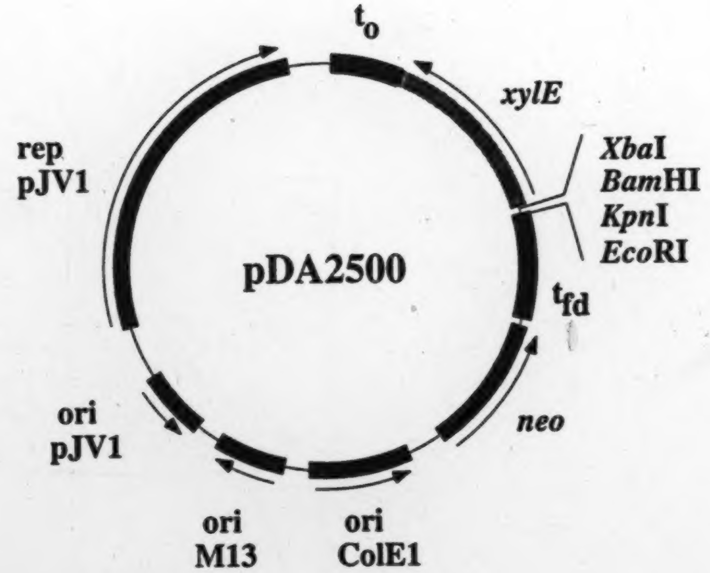

pSET152
PstI
pIJ2925
PstI
pDA311
Bg P
att
P X B K E Bg
EcoRI and
XbaI
Klenow
pDA312
E]/[X
Bg P
att
P Bg
BglII
H
B
[X]
P
Sp
pDA307
Sa
Sa
Dyl11
SalI
H
B
[X]
P
Sp
pDA309
Bg
pDA305
BglII and
HindIII
BamHI and
HindIII
pDA313
Bg
PSp
H
t_{fd}
E S K B X
xylE
t_o B/Bg
BglII
pDA2000
Bg
E]/[X
P
att
P
Bg
H
t_{fd}
E S K B X
xylE
t_o B/Bg

Figure 20. *E. coli/Streptomyces* promoter probe shuttle vectors pDA2000 and pDA2500. Unique restriction sites within the MCS that can be used for cloning promoter fragments are listed.


t_o
xylE
*Xba*I
*Bam*HI
*Kpn*I
*Sac*I
*Eco*RI
pDA2000
t_{fd}
att
neo
ori
ColE1


t_o
xylE
*Xba*I
*Bam*HI
*Kpn*I
*Eco*RI
rep
pJV1
pDA2500
t_{fd}
neo
ori
pJV1
ori
M13
ori
ColE1

IV.2.1 Verification of the pDA2000-based plasmid integration by genomic PCR

Since the pDA2000-based plasmids contain the *attP* site and integrate into the chromosome of *S. clavuligerus* a simple isolation of plasmid DNA could not be used to confirm presence of the plasmid or check the stability of the plasmid and the insert. Instead the genomic DNA (gDNA) was purified from each strain for analysis by PCR using the primers Dyl21 and Dyl33 (Fig. 21A). In this procedure the MCS and the insert of the integrated plasmid were selectively amplified from the gDNA. The Dyl21 primer was constructed using known pIJ4083 MCS sequence (47), while Dyl33 was constructed from sequence information gained by sequencing pDA302. If desired these oligonucleotide primers could also be used as primers for sequencing the insert to verify that the correct promoter fragment was cloned.

Fig. 21B demonstrated that insertion of the pDA2000-based vector within the chromosomal DNA could be confirmed by PCR of gDNA rather than using Southern blot analysis. The MCS/insert from pDA2000, pDA2001 (*lat* promoter) and pDA2003 (*lat* gene) were specifically amplified from the gDNA of the strains created by transformation with the pDA2000-based integration vectors. The amplified PCR product from gDNA was compared to the product amplified from free plasmid isolated from *E. coli*. In all cases the size of the product amplified from the gDNA and the free plasmid were identical. The amplification of the MCS/insert however does not prove plasmid integration into the att site but successful amplification does prove integration of the pDA2000-based vector into the chromosome.

IV.3 Promoter probe analysis of the *lat* gene

Promoter probe analysis was used to examine the activity of the putative *pcbAB* promoter region within the *lat* coding region and the stem-loop structure located in the *lat*-*pcbAB* intergenic region. Fragments containing increasing amounts of the *lat* gene were inserted into the single copy promoter probe vector pDA2000 to examine the promoter strength of each fragment within wildtype *S. clavuligerus*.

IV.3.1 Construction of pDA2000-based promoter probe plasmids with *lat* gene fragments

Fragments containing increasing amounts of the *lat* gene were inserted into pDA2000 (Fig. 22). The 225 bp *Eco*RI/*Nco*I fragment would contain the *lat* promoter as the *lat* start codon was within the *Nco*I site. This construct (pDA2001) would provide a measure of the strength of the *lat* promoter. The 1.1 kb *Eco*RI/*Kpn*I fragment used to

Figure 21. Confirmation of pDA2000-based vector insertion by genomic PCR of transformed strains. **A.** Sequence of the pDA2000 MCS. The oligonucleotide primers Dyl21 and Dyl33 allow for the specific amplification of the MCS plus the cloned insert from the pDA2000-based constructs. Unique restriction sites present within the MCS are noted. **B.** Aliquots of each PCR reaction amplified from the following templates were separated on a 1.5% (w/v) agarose gel. Lane 1. pDA2000 plasmid purified from *E. coli*, 2. genomic DNA from pDA2000 transformed strain, 3. genomic DNA from pDA2000 transformed strain,, 4. pDA2001 plasmid containing the *lat* promoter purified from *E. coli*, 5. genomic DNA from pDA2001 transformed strain, 6. genomic DNA from pDA2001 transformed strain, 7. pDA2003 plasmid containing the entire *lat* gene purified from *E. coli*, 8. genomic DNA from pDA2003 transformed strain, and 9. genomic DNA from pDA2003 transformed strain.

A.

Dyl33

```
aattcatttactaacgtttggaaagacgacaaaactttagatctgcagccaagcttatcgatgataagctgtcaaacatga
ttaagtaaatgattgcaaacctttctgctgttttgaaatctagacgtcggttcgattagctactattcgacagtttgtact
```

*Eco*RI *Sac*I *Kpn*I *Bam*HI *Xba*I

```
gaattcgagctcggtacccggggatcctctagagtcactttcgccacgttggcggaaacaaacctgacaacatgaactatg
cttaagctcgagccatgggcccctaggagatctcagtgaaagcggtgcaaccgcctttgtttggactgttgtacttgatac
```

Dyl21

B.

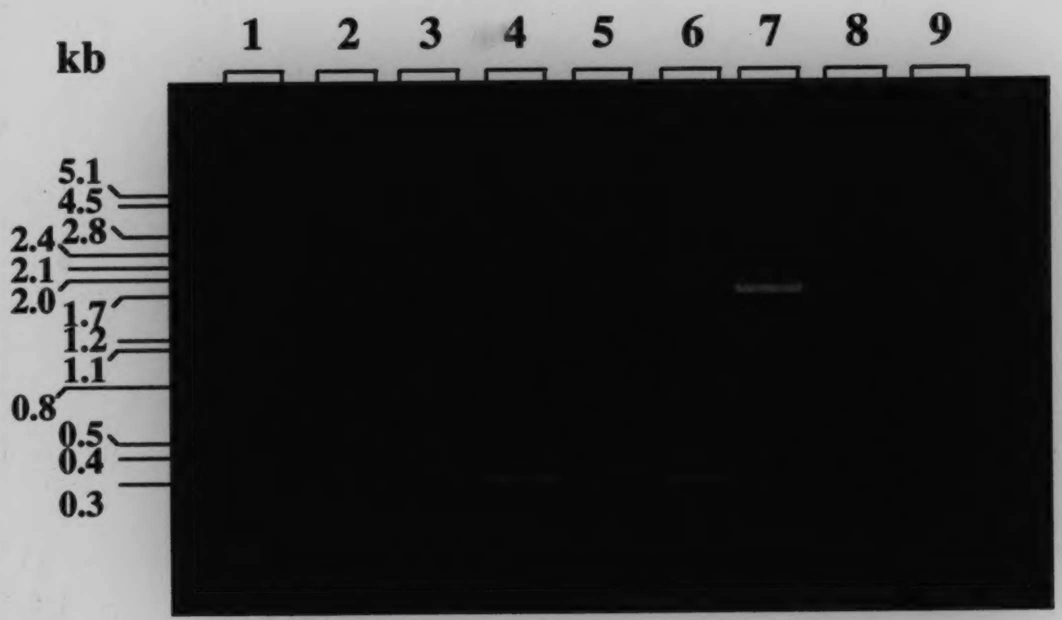
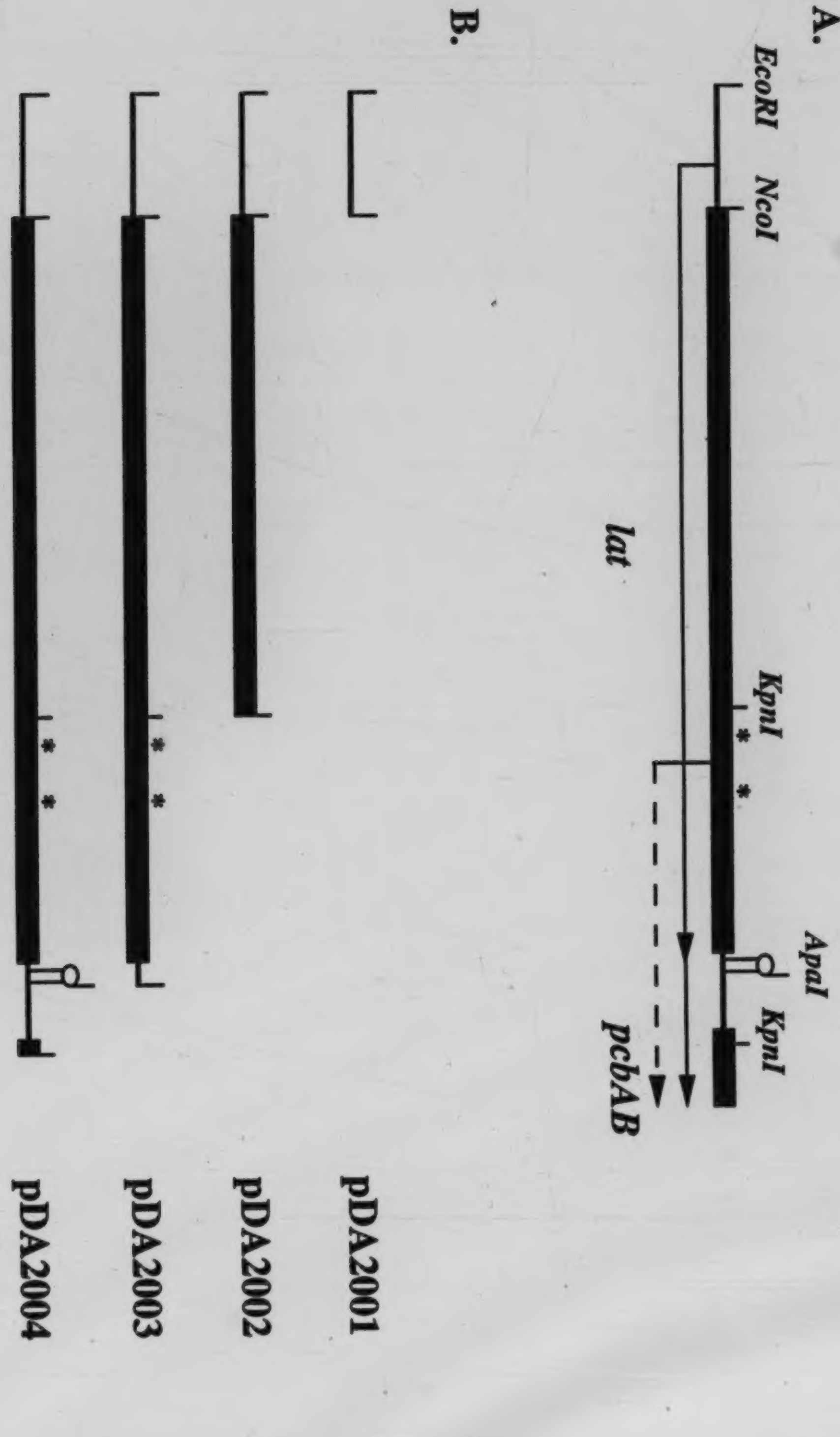


Figure 22. Map of the *lat* gene-containing DNA and pDA2000 constructs created with fragments of DNA from this region. **A.** Restriction map of the *lat-pcbAB* region to be characterized by promoter probe analysis. Coding regions of the genes are indicated by closed boxes and single lines represent the intergenic regions. Solid arrows represent transcription initiating at the *lat* promoter and the dashed arrow represents possible transcription initiating at the putative *pcbAB* promoter. The asterisks mark the location of the putative *pcbAB* promoter region while the hairpin represents the stem-loop structure. **B.** Diagram of the *lat*-containing DNA fragments inserted into the promoter probe vector pDA2000. The pDA2001 vector contains the *lat* promoter, the pDA2002 vector contains the *lat* promoter and some of the *lat* coding region, the pDA2003 vector contains the entire *lat* gene and the pDA2004 vector contains the entire *lat* gene and the downstream stem-loop structure.


A.
EcoRI
NcoI
KpnI
ApaI
KpnI
lat
pcbAB
B.
pDA2001
pDA2002
pDA2003
pDA2004

create pDA2002 includes a greater amount of the *lat* gene but still excludes the putative *pcbAB* promoter region. The 1.6 kb *Eco*RI/*Apa*I fragment was used to create pDA2003, this fragment contains the putative *pcbAB* promoter and part of the downstream stem-loop structure. However, the stem-loop structure was not expected to be able to function as a transcription terminator because of the *Apa*I site within the stem-loop. The 1.8 kb *Eco*RI/*Kpn*I/*Kpn*I fragment was used to create pDA2004 and this fragment would include the putative *pcbAB* promoter and the intact downstream stem-loop structure. Wildtype *S. clavuligerus* protoplasts were transformed with the four different *lat* gene-containing promoter probe constructs as well as the pDA2000 control vector.

IV.3.2 Promoter probe analysis of *lat* gene constructs in wildtype *S. clavuligerus*

Two independent strains generated from each transformation were grown for 48 hours in trypticase soy broth plus 1% (w/v) starch (TSBS). The harvested cell pellets were resuspended in phosphate breakage buffer (see. II.2.4.5), disrupted by sonication and the cell-extracts were then assayed for C23O activity (see II.2.4.5.1). Average specific activities were obtained from cell-extracts generated from two separate fermentations.

Low levels of C23O activity were measured from *S. clavuligerus* transformed with the pDA2000 control plasmid (Table 11). Much higher levels of C23O activity were found in strains transformed with the constructs containing the *lat* gene fragments. A decrease in C23O specific activity occurred in constructs containing longer *lat* gene fragments. No dramatic difference was observed for the constructs which contain the putative *pcbAB* promoter region. As well, minimal differences in activity were detected for strains containing the pDA2003 or pDA2004 plasmids which differ by whether or not they contain a region predicted to from a functional stem-loop structure in the mRNA between the *lat* and *pcbAB* coding regions. Comparison of C23O specific activities of pDA2001 through pDA2004 suggests that neither the putative *pcbAB* promoter nor the *lat* terminator are functional under the conditions examined. The near 2-fold decrease in C23O specific activity by the pDA2002 cultures when compared to the pDA2001 cultures both of which contain the *lat* promoter suggests that promoter probe analysis may be providing at best a semiquantitative measurement of the promoter strength of each fragment. Therefore promoter probe analysis data should be combined with data generated by transcript analysis or characterization of mutants prepared by homologous recombination before strong conclusions are drawn.

Table 11. **Promoter probe analysis of *lat* gene fragments**

	Catechol dioxygenase specific activity[a]			
plasmid	strain A[b]	strain B[b]	average	average[c] / pDA2001
pDA2000	6.1	7.3	6.7	5.2
pDA2001	123	137	130	100
pDA2002	68	67	68	52.3
pDA2003	80	98	89	68.5
pDA2004	64	66	65	50.0

[a] C23O specific activity values are given as EU x 10^{-6}/mg protein.

[b] Specific activities values are an average of two separate fermentations for each strain. Strains A and B are two independent transformants generated during the same wildtype *S. clavuligerus* transformation.

[c] Values are calculated as a percentage of the specific activity obtained for the pDA2001 plasmid.

IV.4 Mutation of the *lat* gene by gene replacement

Results of promoter probe plasmid studies using the single copy vector are consistent with transcriptional studies done by S1 analysis. S1 nuclease transcript analysis provided no evidence of transcription originating from the putative *pcbAB* promoter within *lat* and no evidence of transcription termination at the putative terminator located between *lat* and *pcbAB* (169, 175). Failure to detect increased promoter activity in fragments containing the putative *pcbAB* promoter suggests that this promoter region does not support detectable levels of transcription when present at single copy. As well a major decrease in C23O specific activity did not occur when the *lat* stem-loop structure was intact; suggesting that the stem-loop structure downstream of the *lat* coding region would not prevent transcription from proceeding into the downstream genes. From these results it follows that transcription of *pcbAB* can only occur as part of a polycistronic transcript initiating at the *lat* promoter. However, constructs containing the putative *pcbAB* promoter did show some minor increases in C23O activity levels and constructs containing the putative terminator did show somewhat decreased levels of C23O activity, therefore some uncertainty remained.

To provide more definitive proof of this transcriptional relationship, a series of *lat* mutants were prepared by homologous recombination. The goal of Jensen and co-workers in creating the *lat*::*apr* mutant (111) was to demonstrate that interruption of transcription from the *lat* promoter by insertion of the *apr* marker would eliminate ACVS activity if the putative *pcbAB* promoter was not functional. The *apr* marker was inserted into the unique *Kpn*I site within the center of the *lat* coding region and with an orientation opposite to the *lat* gene. The *apr* marker was inserted upstream of the putative *pcbAB* promoter region and therefore would not alter transcription originating from within this region. The *lat*::*apr* mutant however still retained ACVS activity but S1 nuclease transcript analysis could not confirm that the putative *pcbAB* promoter was responsible for this ACVS activity. Therefore, three additional sets of *lat* mutants were generated by homologous recombination which would unequivocally determine whether the putative *pcbAB* promoter within the *lat* coding region was capable of initiating *pcbAB* transcription and answer many of the questions about the transcriptional organization of the *lat*, *pcbAB* and *pcbC* genes.

The first set of *lat* mutants (Δ*lat*::*tsr*) resulted from the replacement of part of the *lat* coding region with the *tsr* marker. The Δ*lat*::*tsr* mutants are similar to the *lat*::*apr* strain except that the resistance marker has been changed to *tsr* which allows the strain to be complemented with pSET152-based plasmids and part of the *lat* coding region has been deleted. The second set of *lat* mutants (Δ*lat*::*tsr*/term) resulted from the replacement of the *lat* promoter and part of the *lat* coding region with the *tsr* marker followed by a terminator. In both of the *lat* partial deletion mutant sets the putative *pcbAB* promoter region was still present but any transcriptional readthrough from upstream promoter sequences would be interrupted or terminated, respectively (Fig. 23). The third set of *lat* mutants (Δ*lat*) resulted from replacement of the mutated *lat* gene containing the *apr* marker within the *lat*::*apr* mutant with a mutated *lat* gene which contains a deletion removing the *Eco*47III/*Kpn*I fragment of the *lat* gene. These mutants would be unable to make a functional LAT protein but transcription of the downstream genes in the polycistronic transcript should not be altered because no resistance marker remains. Again, the putative *pcbAB* promoter region was present in the Δ*lat* mutants.

The mutants were generated by a procedure inviting homologous recombination between the targeting vector carrying the mutagenic gene and the parental chromosome (see II.2.5). The targeting vector crosses over into the chromosome via the homologous DNA insert and then crosses back out to complete the double crossover. When the single crossovers occur on either side of the mutated gene within the targeting vector the mutant gene would replace the wildtype gene in the chromosome. Putative mutants were selected for maintenance of the thiostrepton resistant phenotype (tsr^R; deletion marker) but loss of the hygromycin resistant phenotype (hyg^R; shuttle vector marker) phenotype. In the situation where the *lat*::*apr* mutant strain was "repaired" to remove the *apr* marker the putative mutants were screened for the loss of the apramycin resistant phenotype (apr^R; disruption marker) and loss of the tsr^R phenotype (shuttle vector marker).

IV.4.1 Partial deletion of *lat* and replacement with the *tsr* marker (Δ*lat*::*tsr* mutants)

The first set of *lat* mutants created resulted from the replacement of the *Nco*I/*Eco*47III fragment within *lat* coding region with the *tsr* marker (Fig. 24). The plasmid constructs pDA561 and pDA562 carried a contiguous stretch of *S. clavuligerus* DNA which extends from *orf11* to *lat* except that the start of the *lat* coding region was removed and replaced with the *tsr* marker in both orientations. Each plasmid construct was introduced into protoplasts of wildtype *S. clavuligerus* by transformation and two transformants were allowed to sporulate in the absence of antibiotic selection to promote the loss of free plasmid. Putative mutants were selected for maintenance of the tsr^R phenotype but loss of the hyg^R phenotype. Putative mutants were not immediately obtained from transformants containing the pDA562 plasmid and therefore four mutants from the pDA561 transformants were characterized. The pDA561 plasmid had the *tsr* marker in the opposite orientation to the *lat* gene (shown in Fig. 24) and since the object

Figure 23. Mutated *lat* genes generated by partial deletion of the *lat* gene via homologous recombination. The dark shaded boxes represent the adjacent *pcbAB* gene and the lighter shaded boxes represent the *lat* gene while the clear and hatched boxes represent the *tsr* marker and the FKMT terminator, respectively. The asterisks mark the location of the putative *pcbAB* promoter region and the dashed arrow represents possible transcription originating at the *pcbAB* promoter. The thickness of the solid arrows represent the expected level of transcription originating from the *lat* promoter in each *lat* mutant.

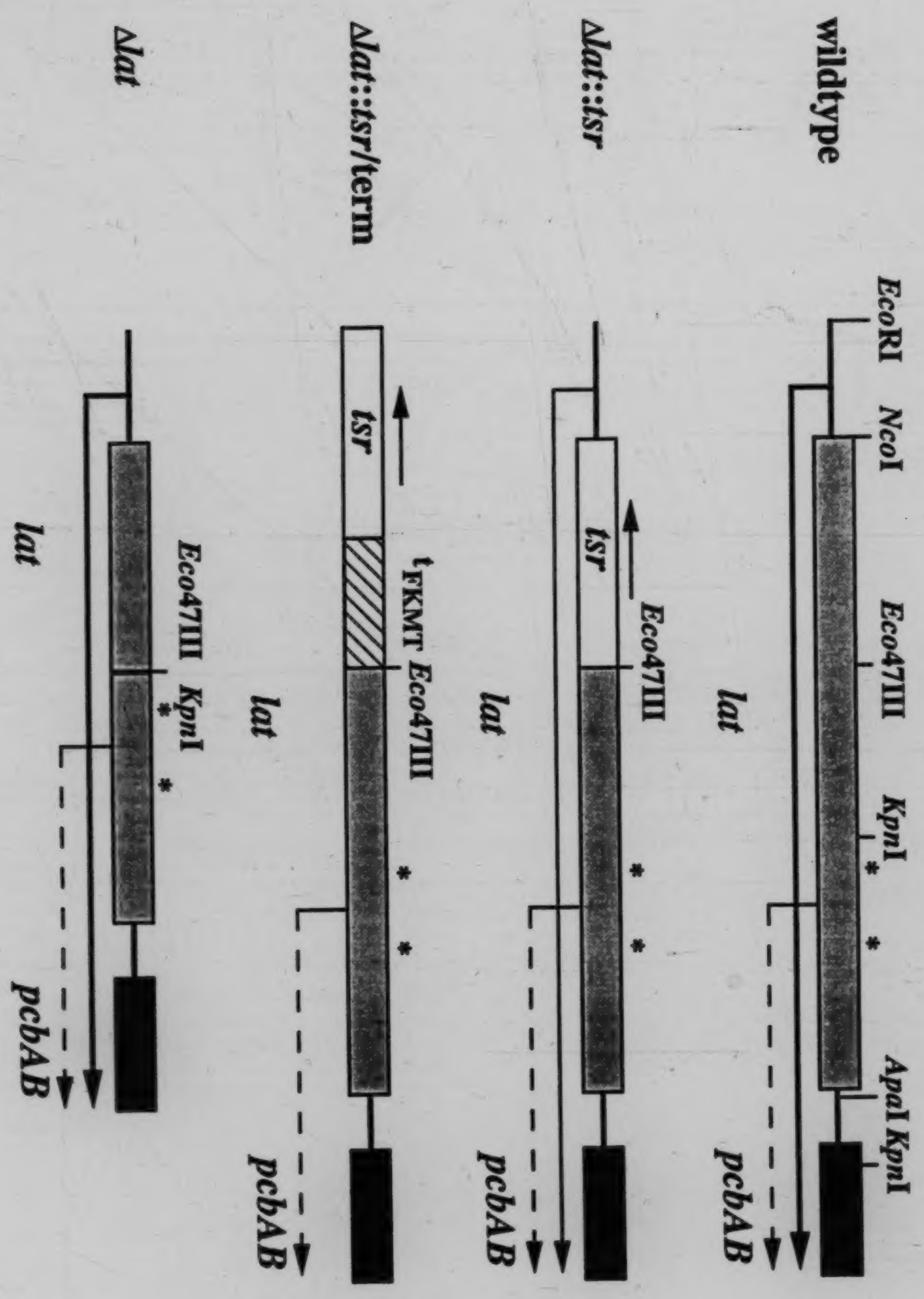
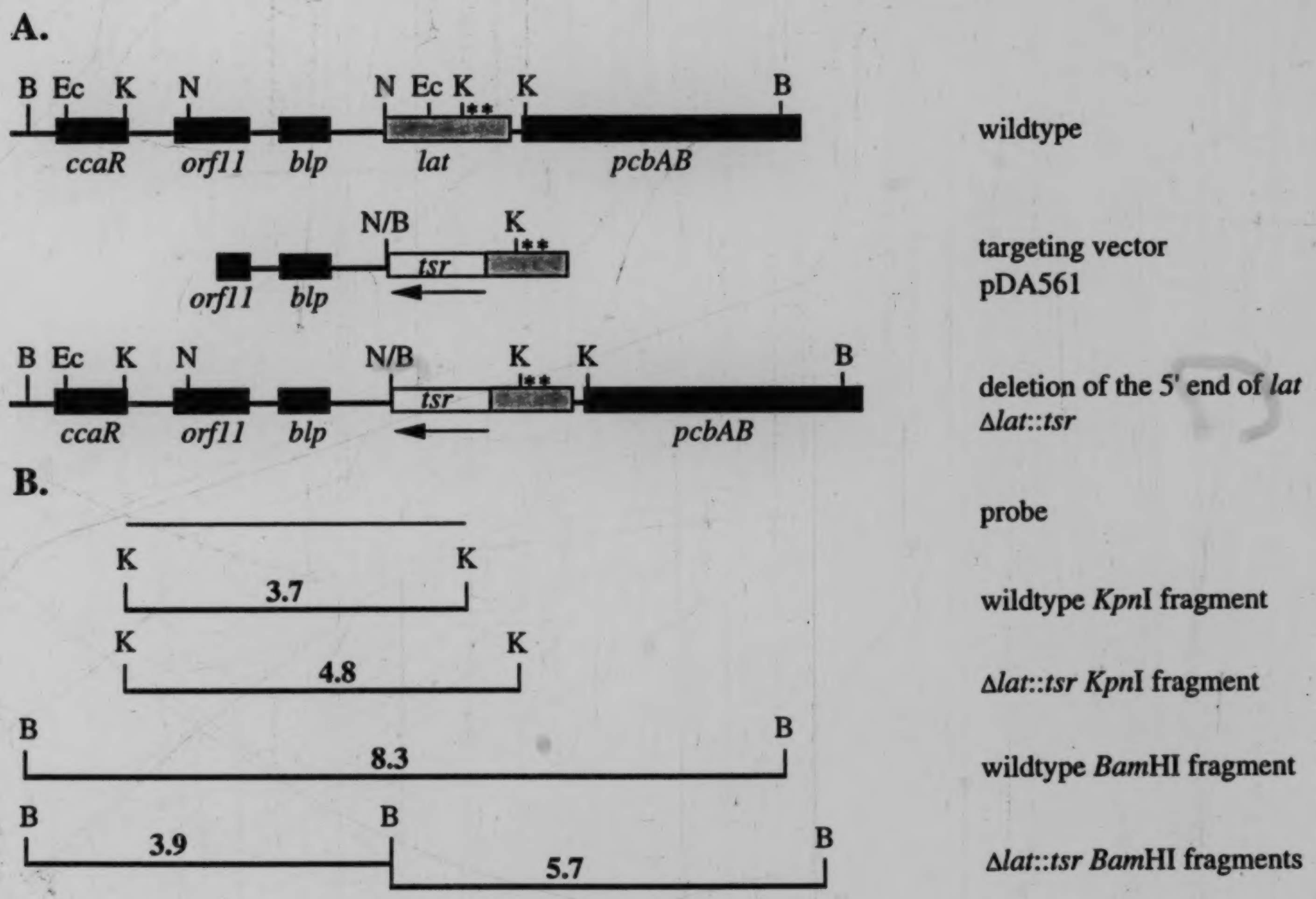

wildtype
EcoRI
NcoI
Eco47III
KpnI
ApaI
KpnI
lat
pcbAB
Δlat::tsr
tsr
Eco47III
lat
pcbAB
Δlat::tsr/term
tsr
tFKMT
Eco47III
lat
pcbAB
Δlat
Eco47III
KpnI
lat
pcbAB

Figure 24. Strategy for the partial deletion of the *lat* gene and replacement with *tsr*. **A.** The dark shaded boxes represent ORFs adjacent to *lat* in the chromosome and the lighter shaded boxes represent the *lat* gene within the wildtype cephamycin cluster. The asterisks mark the location of the putative *pcbAB* promoter region. The targeting vector pDA561 contained a *S. clavuligerus* DNA fragment in which the *NcoI/Eco*47III fragment from the *lat* coding region was deleted and replaced with the *tsr* marker, represented by the clear box. The mutant chromosome results as the partially deleted *lat* gene containing the *tsr* marker replaces the wildtype *lat* gene in the chromosome. **B.** Southern hybridization pattern of *Kpn*I or *Bam*HI digested genomic DNA from wildtype and Δ*lat*::*tsr* mutants using the 3.7 kb *Kpn*I fragment as a probe. Abbreviations for restrictions sites: B for *Bam*HI, Ec for *Eco*47III, K for *Kpn*I and N for *Nco*I.


A.
B
Ec
K
N
N
Ec
K
**
K
B
ccaR
orf11
blp
lat
pcbAB
wildtype
N/B
K
**
tsr
orf11
blp
targeting vector
pDA561
B
Ec
K
N
N/B
K
**
K
B
tsr
ccaR
orf11
blp
pcbAB
deletion of the 5' end of *lat*
Δ*lat::tsr*
B.
probe
K
K
3.7
wildtype *Kpn*I fragment
K
K
4.8
Δ*lat::tsr Kpn*I fragment
B
B
8.3
wildtype *Bam*HI fragment
B
B
B
3.9
5.7
Δ*lat::tsr Bam*HI fragments

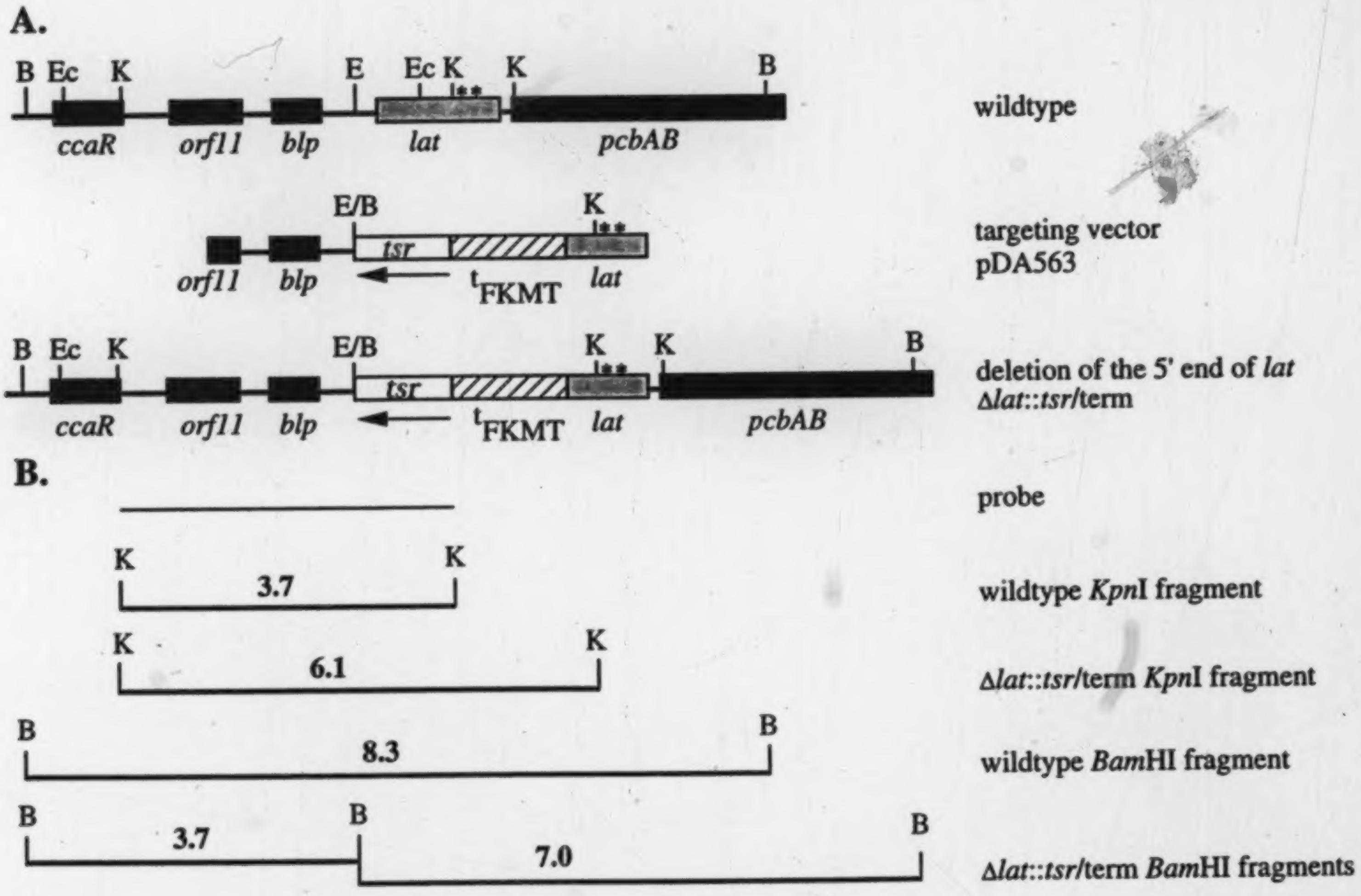
of these mutations was to create a polar mutant no further efforts were made to make the pDA562 orientation.

Southern blot analysis of restriction enzyme digested genomic DNA (gDNA) from parental and mutant strains confirmed the replacement of the target gene in the chromosome with the mutated version. In all cases one of the gDNA digests was designed to overlap between DNA present in the targeting vector and DNA present only on the chromosome as this can confirm that a double crossover had occurred as opposed to a single crossover integrating the targeting vector or the presence of free plasmid in the strain. The digested gDNA was probed with the 3.7 kb *Kpn*I fragment containing parts or all of the *ccaR*, *orf11*, *blp* and *lat* genes (Fig. 24). The wildtype *Kpn*I digested gDNA gave a hybridizing band of 3.7 kb (Fig. 27, lane 1) while the Δ*lat*::*tsr* mutants gave 4.8 kb hybridizing bands (lanes 2-3). The wildtype *Bam*HI digested gDNA gave a hybridizing band of 8.3 kb (lane 10), while the Δ*lat*::*tsr* mutants gave hybridizing bands of 3.9 kb and 5.7 kb (lanes 11-12). The 5.7 kb *Bam*HI band from the Δ*lat*::*tsr* mutants could only be observed after long exposures. The above results were all consistent with a double crossover event having occurred in the Δ*lat*::*tsr* mutants of *S. clavuligerus*. To create a concise figure only two mutants from each *lat* mutant set were included, rather than including a figure for each mutant set generated. Southern analysis characterizing all four individual mutants in a particular set was done previously (data not shown) and in all cases the four mutants strains gave an identical restriction pattern.

IV.4.2 Partial deletion of *lat* and replacement with the *tsr* marker and FKMT terminator (Δ*lat*::*tsr*/term mutants)

The second set of *lat* mutants created resulted from the replacement of the *Eco*RI/*Eco*47III fragment containing the *lat* promoter and a portion of *lat* coding region with the *tsr* marker followed by the 31-demethyl-FK506-*o*-methyltransferase (FKMT) terminator (Fig. 25). The plasmid constructs pDA563 and pDA564 carried a contiguous stretch of *S. clavuligerus* DNA which extends from *orf11* to *lat* except that the *lat* promoter and the start of the *lat* coding region were removed and replaced with the *tsr* marker in both orientations followed by the FKMT terminator. Each plasmid construct was introduced into protoplasts of wildtype *S. clavuligerus* by transformation and two transformants were allowed to sporulate in the absence of antibiotic selection to promote the loss of free plasmid. Putative mutants were selected for maintenance of the tsr^R phenotype but loss of the hyg^R phenotype. Putative mutants were not obtained from transformants containing the pDA564 plasmid and therefore four mutants from the

Figure 25. Strategy for the partial deletion of the *lat* gene and replacement with *tsr* and the FKMT terminator. **A.** The dark shaded boxes represent ORFs adjacent to *lat* in the chromosome and the lighter shaded boxes represent the *lat* gene within the wildtype cephamycin cluster. The asterisks mark the location of the putative *pcbAB* promoter region. The targeting vector pDA563 contained a *S. clavuligerus* DNA fragment in which the *Eco*RI/*Eco*47III fragment from the *lat* coding region was deleted and replaced with the *tsr* marker and the FKMT terminator, represented by the clear box and the hatched box, respectively. The mutant chromosome results as the partially deleted *lat* gene containing the *tsr* marker and FKMT terminator replaces the wildtype *lat* gene in the chromosome. **B.** Southern hybridization pattern of *Kpn*I or *Bam*HI digested genomic DNA from wildtype and Δ*lat*::*tsr*/term mutants using the 3.7 kb *Kpn*I fragment as a probe. Abbreviations for restrictions sites: B for *Bam*HI, E for *Eco*RI, Ec for *Eco*47III and K for *Kpn*I.


A.
B Ec K
E
Ec K
K
B
ccaR
orf11
blp
lat
pcbAB
wildtype
E/B
K
tsr
orf11
blp
tFKMT
lat
targeting vector
pDA563
B Ec K
E/B
K
K
B
tsr
ccaR
orf11
blp
tFKMT
lat
pcbAB
deletion of the 5' end of lat
Δlat::tsr/term
B.
probe
K
K
3.7
wildtype KpnI fragment
K
K
6.1
Δlat::tsr/term KpnI fragment
B
B
8.3
wildtype BamHI fragment
B
B
B
3.7
7.0
Δlat::tsr/term BamHI fragments

pDA563 transformants were characterized. The pDA563 plasmid has the *tsr* marker in the opposite orientation (shown in Fig. 25) to the *lat* gene.

Again the mutation was confirmed by Southern blot analysis. The wildtype *Kpn*I digested gDNA was probed with the 3.7 kb *Kpn*I fragment (Fig. 25). The wildtype gDNA gave a hybridizing band of 3.7 kb (Fig. 27, lane 1) while the Δ*lat*::*tsr*/term mutants gave 6.1 kb hybridizing bands (lanes 4-5). The wildtype *Bam*HI digested gDNA gave a hybridizing band of 8.3 kb (lane 10), while the Δ*lat*::*tsr*/term mutants gave hybridizing bands of 3.7 kb and 7.0 kb (lanes 13-14). The 7.0 kb *Bam*HI band from the Δ*lat*::*tsr*/term mutants was very faint in longer exposures. The above results were all consistent with a double crossover event having occurred in the Δ*lat*::*tsr*/term mutants of *S. clavuligerus*.

IV.4.2.1 Construction of the FKMT-containing promoter probe plasmids

The FKMT terminator was removed from the pHM8a *Streptomyces* expression vector (153) for insertion into the *lat* gene to prevent transcription originating at upstream promoter sequences from proceeding into the *pcbAB* and *pcbC* genes. The *tsr* marker was placed upstream of the terminator and used as a phenotypic marker for the double crossover. The FKMT terminator was inserted behind the *tsr* marker which was oriented in both directions and therefore in one orientation it would transcribe directly into the terminator and in the other orientation it would transcribe away from the terminator.

To test that the terminator was capable of preventing transcriptional readthrough from the *tsr* marker, both of the *tsr* marker and terminator fragments were inserted in front of the partial *xylE* reporter gene cassette in pDA305. The resulting fragments containing the *tsr* marker, the terminator and the partial reporter gene cassette were then inserted into pSET152. The constructs were made in pSET152 rather than pDA2000 because initially it was thought that the FKMT terminator in pHM8a was actually the fd terminator and this would then be duplicated in a pDA2000-based plasmid. To test the effectiveness of the FKMT terminator, these two constructs (pDA2007 and pDA2008) were transformed into wildtype *S. clavuligerus* protoplasts and compared to the pDA2000 negative control plasmid.

IV.4.2.2 Promoter probe analysis of the effectiveness of the FKMT terminator

Two independent strains generated from each transformation were grown for 48 hours in TSBS. The harvested cell pellets were resuspended in phosphate breakage buffer, disrupted by sonication and the cell-extracts were then assayed for C23O activity. Average specific activities were obtained from cell-extracts generated from two separate fermentations.

The pDA2007 and pDA2008 constructs both showed very weak C23O specific activity levels of 8 EU x 10^{-6}/mg of protein. These C23O specific activity values were comparable to the pDA2000 vector without an insert which had a specific activity of 11 EU x 10^{-6}/mg of protein. Therefore it was felt that the FKMT terminator was successfully preventing transcriptional readthrough from upstream promoter sequences.

IV.4.3 Partial deletion of *lat* after removal of the *apr* marker from *lat*::*apr* (Δ*lat* mutants)

The third set of *lat* mutants were designed to create a mutant strain in which no functional LAT protein was made but transcription of the downstream genes in the polycistronic message was not altered (Fig. 26). The plasmid constructs pDA565 and pDA566 carried a contiguous stretch of *S. clavuligerus* DNA which extends from *orf11* to *pcbAB* except that either a frameshift was introduced at the unique *Kpn*I site in the *lat* coding region or a deletion removing the *Eco*47III/*Kpn*I fragment internal to the *lat* coding region was introduced. Each plasmid construct was introduced into protoplasts of the *lat*::*apr* mutant strain (111) by transformation and then two transformants were allowed to sporulate in the absence of antibiotic selection to promote the loss of free plasmid. Putative mutants were selected for the loss of the apr^R and tsr^R phenotype. The loss of the apr^R phenotype from the *lat*::*apr* strain would demonstrate that the mutant *lat* gene in pDA565 or pDA566 had replaced the mutated *lat* gene containing the *apr* marker in the chromosome. Putative mutants were not obtained from transformants containing the pDA565 plasmid and therefore four mutants from the pDA566 transformants (shown in Fig. 26) were characterized.

The mutation was confirmed by Southern blot analysis but the Δ*lat* mutants were compared to the *lat*::*apr* mutant strain since the Δ*lat* mutants were generated from this background. The wildtype *Kpn*I digested gDNA was probed with the 3.7 kb *Kpn*I fragment (Fig. 26). The *lat*::*apr* strain gDNA gave a hybridizing band of 3.7 kb (Fig. 27, lane 6) while the Δ*lat* mutants gave 4.0 kb hybridizing bands (lanes 7-8). The *lat*::*apr* strain *Bam*HI digested gDNA gave a hybridizing band of 6.3 kb (lane 15), while the Δ*lat* mutants gave hybridizing bands of 8.0 kb (lane 16-17). The above results were all consistent with a double crossover event having occurred in the Δ*lat* mutants of *S. clavuligerus*.

Figure 26. Strategy for the deletion of an internal fragment of the *lat* gene. **A.** The dark shaded boxes represent ORFs adjacent to *lat* in the chromosome, the lighter shaded boxes represent the *lat* gene and the clear box represents the *apr* marker within the *lat*::*apr* mutant. The asterisks mark the location of the putative *pcbAB* promoter region. The targeting vector pDA566 contained a *S. clavuligerus* DNA fragment in which the *Eco*47III/*Kpn*I fragment inside the *lat* coding region was deleted. The mutant chromosome results as the partially deleted *lat* gene replaces the mutant *lat* gene containing the *apr* marker in the chromosome. **B.** Southern hybridization pattern of *Kpn*I or *Bam*HI digested genomic DNA from wildtype and Δ*lat* mutants using the 3.7 kb *Kpn*I fragment as a probe. Abbreviations for restrictions sites: B for *Bam*HI, Ec for *Eco*47III and K for *Kpn*I.

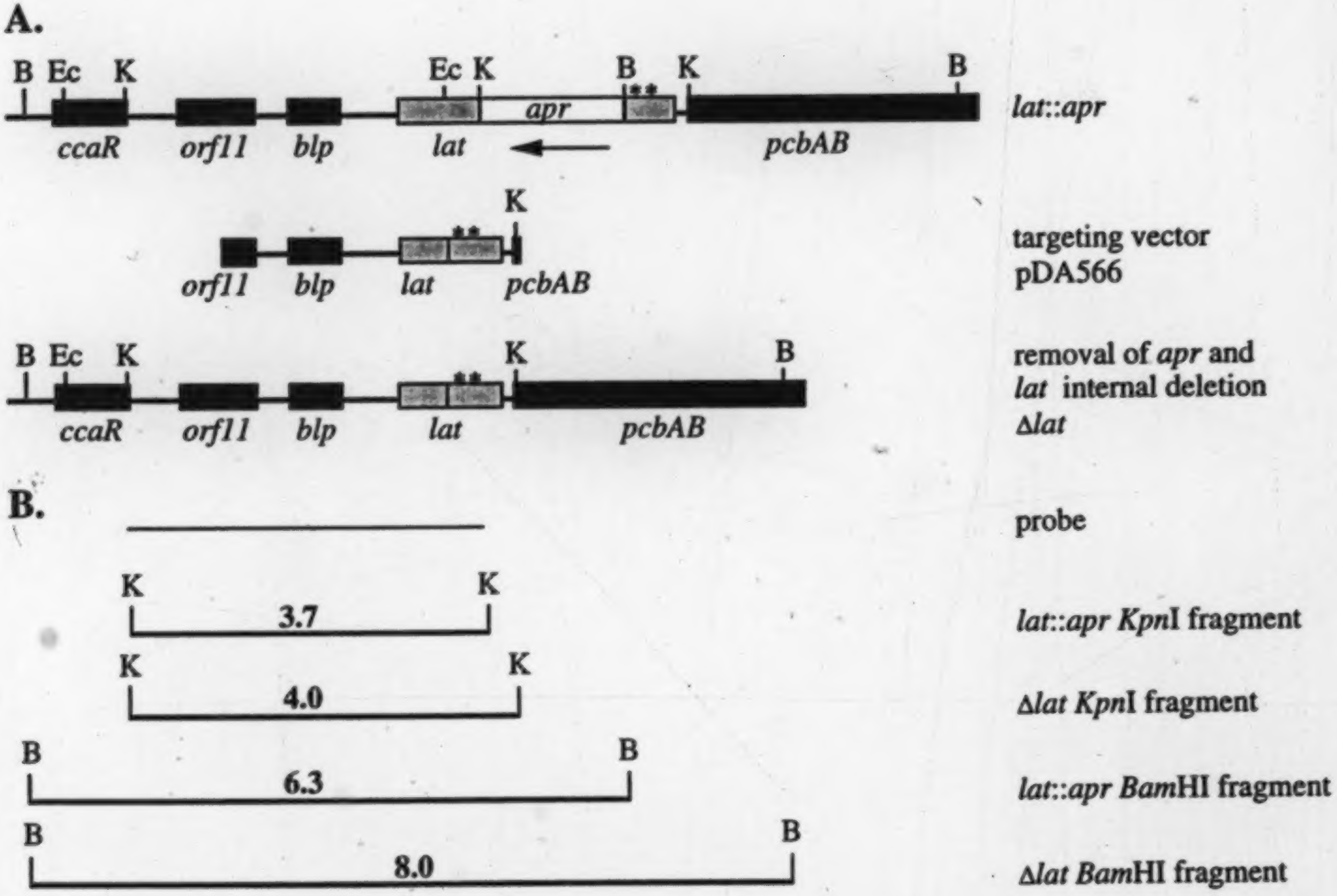
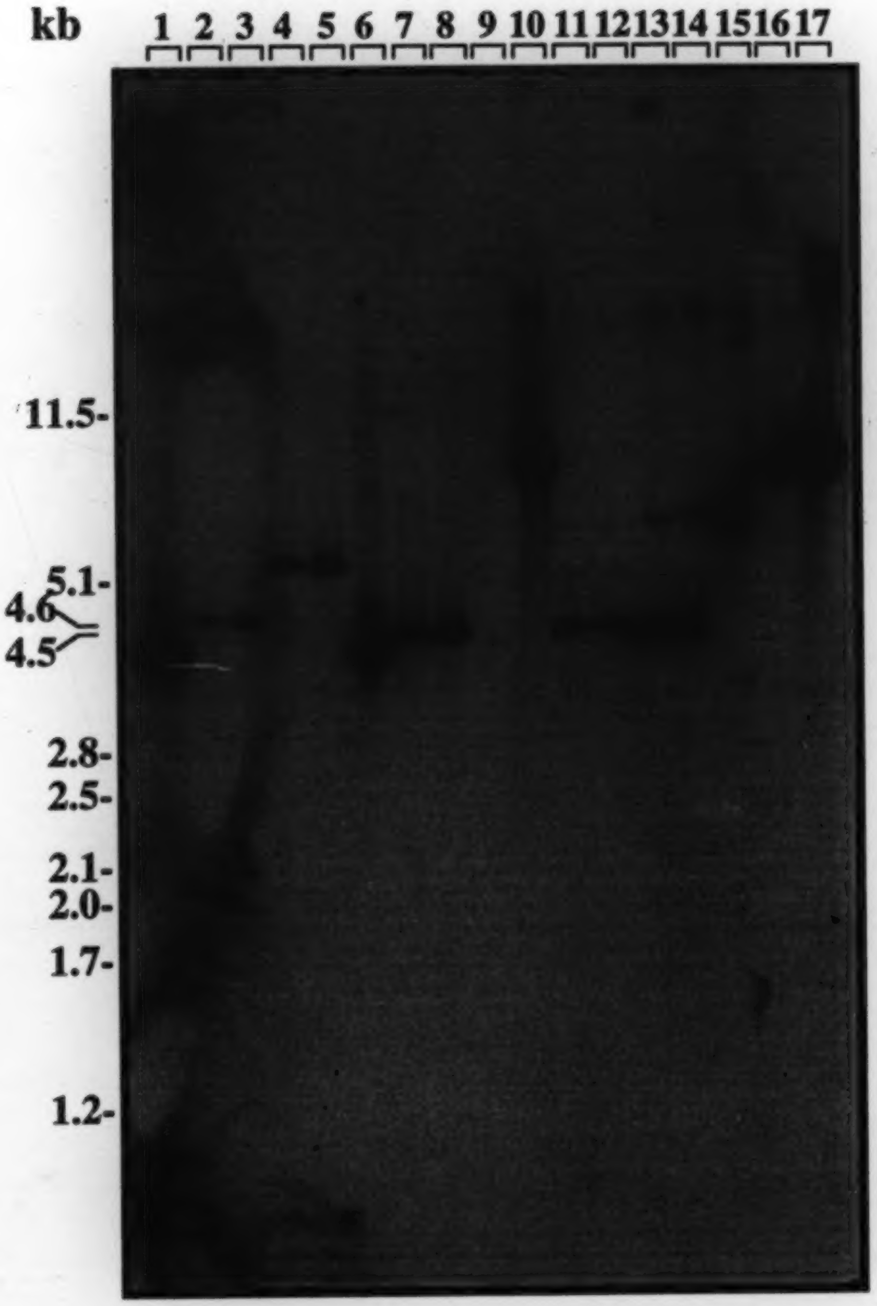

A.
B
Ec
K
Ec
K
B
K
B
apr
lat::apr
ccaR
orf11
blp
lat
pcbAB
K
orf11
blp
lat
pcbAB
targeting vector
pDA566
B
Ec
K
K
B
ccaR
orf11
blp
lat
pcbAB
removal of apr and
lat internal deletion
Δlat
B.
probe
K
K
3.7
lat::apr KpnI fragment
K
K
4.0
Δlat KpnI fragment
B
B
6.3
lat::apr BamHI fragment
B
B
8.0
Δlat BamHI fragment

Figure 27. Confirmation of gene replacement in *lat* mutants by Southern blot analysis. The *Kpn*I or *Bam*HI digested genomic DNA isolated from wildtype and *lat* mutants was hybridized with the 3.7 kb *Kpn*I fragment (see Fig. 24, 25 and 26). *Kpn*I digested genomic DNA (Lanes 1. wildtype *S. clavuligerus*, 2. Δ*lat*::*tsr*A1 mutant, 3. Δ*lat*::*tsr*A2 mutant, 4. Δ*lat*::*tsr*/termA1 mutant, 5. Δ*lat*::*tsr*/termA2 mutant, 6. *lat*::*apr* mutant, 7. Δ*lat*A1 mutant and 8. Δ*lat*A2 mutant), lane 9. size marker and *Bam*HI digested genomic DNA (Lanes 10. wildtype *S. clavuligerus*, 11. Δ*lat*::*tsr*A1 mutant, 12. Δ*lat*::*tsr*A2 mutant, 13. Δ*lat*::*tsr*/termA1 mutant, 14. Δ*lat*::*tsr*/termA2 mutant, 15. *lat*::*apr* mutant, 16. Δ*lat*A1 mutant, and 17. Δ*lat*A2 mutant) were separated on a 0.7% (w/v) agarose gel.


kb
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
11.5-
5.1-
4.6
4.5
2.8-
2.5-
2.1-
2.0-
1.7-
1.2-

IV.4.4 Additional Southern blot analyses of *lat* mutant strains

The Southern blots used to confirm the nature of the mutant strains were stripped and reprobed with other fragments of DNA such as the *tsr* marker. The results from these secondary hybridizations are only discussed in the text and reproductions of the autoradiographs are not included.

When the *Kpn*I digested gDNA was hybridized with the *tsr* marker the Δ*lat*::*tsr* mutants gave a 4.8 kb hybridizing band and the Δ*lat*::*tsr*/term mutants gave a 6.1 kb hybridizing band while the there was no hybridization by the wildtype, Δ*lat* or *lat*::*apr* gDNA lanes. This confirmed the presence of the *tsr* marker within the Δ*lat*::*tsr* mutants and the Δ*lat*::*tsr*/term mutants.

When the *Kpn*I digested gDNA was hybridized with the *apr* marker the *lat*::*apr* mutant gave a 2.0 kb hybridizing band while the wildtype, Δ*lat*, Δ*lat*::*tsr* and the Δ*lat*::*tsr*/term mutants gDNA gave no hybridization. This confirmed the absence of the *apr* marker within the Δ*lat* mutants.

When the *Kpn*I digested gDNA was hybridized with a fragment of DNA containing the *pcbC* gene all of the strains gave hybridizing bands of 2.7 kb and 0.5 kb. This confirmed the presence of the *pcbC* gene within the Δ*lat*::*tsr*/term mutants.

The *Kpn*I digested gDNA of all the strains did not contain any DNA sequences that hybridized with pJOE829 demonstrating double crossover and removal of the targeting vector.

IV.5 Characterization of *lat* mutants created by homologous recombination

The three different sets of *lat* mutants generated by homologous recombination were designed to determine whether the putative promoter within the *lat* coding region was responsible for *pcbAB* transcription and whether the *lat* terminator could prevent transcription into the downstream genes.

The Δ*lat*::*tsr* mutants have the replacement of part of the *lat* coding region with the *tsr* marker but the putative *pcbAB* promoter region was still present. The Δ*lat*::*tsr*/term mutants have the replacement of the *lat* promoter and part of the *lat* coding region with the *tsr* marker followed by the FKMT terminator. Again the putative *pcbAB* promoter region was still present but any transcriptional readthrough from upstream promoter sequences would be eliminated by the transcriptional terminator. The Δ*lat* mutants resulted from the deletion of the *Eco*47III/*Kpn*I fragment internal to the *lat* coding region as a result of removing the *apr* marker from the *lat*::*apr* strain. It was assumed that removal of this 354 bp fragment would not greatly affect stability of the transcript. Again the putative *pcbAB* promoter region was present.

IV.5.1 Characterization of cephamycin C production by the *lat* mutants

Four independent strains from each of the three sets of *lat* mutants (Δ*lat*::*tsr*, Δ*lat*::*tsr*/term and Δ*lat*) were grown in TSBS or TSBS plus 2 mM αAA. Half of the cultures were supplemented with αAA, the product of LAT and PCD activities, to compensate for the mutation of the *lat* gene. Samples were taken at both 48 and 72 hours and assayed for the production of cephamycin. Cephamycin production was determined by bioassay using the penicillin supersensitive strain *E. coli* ESS (see II.2.4.4.1.).

The results of cephamycin production during the fermentations are summarized in Table 12. All of the sets of *lat* mutants were unable to make cephamycin in TSBS confirming the mutation of the *lat* gene and the elimination of the essential LAT protein. The addition of 2 mM αAA to the media allowed the Δ*lat*::*tsr* mutants to make only very small amounts of cephamycin, an average of 1.6% and 1.3% of wildtype cephamycin levels after 48 hours and 72 hours growth, respectively. The addition of αAA to cultures of the Δ*lat*::*tsr*/term mutants was unable to stimulate cephamycin production even after 72 hours growth. The Δ*lat* mutants made an average of 18% and 33% of wildtype cephamycin levels after 48 hours and 72 hours growth, respectively. Both of these values were large increases over their parental strain *lat*::*apr* which only made 12% of wildtype cephamycin levels after 48 hours and 15% after 72 hours. The addition of αAA to wildtype cultures had a minimal effect.

The low levels or lack of cephamycin production in the presence of αAA by the Δ*lat*::*tsr* and the Δ*lat*::*tsr*/term mutants suggests that the putative *pcbAB* promoter is not supporting the production of high levels of ACVS and IPNS protein. The absence of *lat* gene disruption allows the Δ*lat* mutants to produce greater amounts of cephamycin than the *lat*::*apr* mutant.

IV.5.2 Western blot analysis of the *lat* mutants

Cell-extracts were prepared from 72 hour cultures for the four mutants from each particular *lat* mutant set, the *lat*::*apr* mutant and wildtype *S. clavuligerus*, as well cell-extracts were prepared from two mutants from each particular *lat* mutant, wildtype *S. clavuligerus* and the *lat*::*apr* mutant grown in TSBS plus 2 mM αAA (see II.2.4.1). The western blots (see II.2.4.3) were developed using antibodies specific for the LAT, IPNS and desacetoxycephalosporin C (DAOCS) biosynthetic enzymes.

Table 12. Production of cephamycin by the *lat* mutants

	Cephamycin production[a]			
strain	48 hours TSBS	48 hours TSBS with αAA[b]	72 hours TSBS	72 hours TSBS with αAA[b]
wildtype	100	57.8	100	110.8
lat::apr	0	12.0	0	14.6
Δ*lat::tsr*1	0	2.2	0	3.1
Δ*lat::tsr*2	0	2.0	0	1.7
Δ*lat::tsr*3	0	1.2	0	0
Δ*lat::tsr*4	0	1.0	0	0
mean Δ*lat::tsr*	0	1.6	0	1.3
Δ*lat::tsr*/term1	0	0	0	0
Δ*lat::tsr*/term2	0	0	0	0
Δ*lat::tsr*/term3	0	0	0	0
Δ*lat::tsr*/term4	0	0	0	0
mean Δ*lat::tsr*/term	0	0	0	0
Δ*lat*1	0	22.4	0	32.9
Δ*lat*2	0	25.2	0	58.4
Δ*lat*3	0	11.1	0	22.7
Δ*lat*4	0	13.5	0	18.1
mean Δ*lat*	0	18.1	0	33.0

[a] Values are calculated as a percentage of the wildtype strain in TSBS at each timepoint, which was given a value of 100%.

[b] Cultures were supplemented with α-aminoadipate to a 2 mM final concentration.

Western blot analysis was preferred as the method to characterize the mutants rather than enzyme assays for several reasons. With western blot analysis, the presence of all the proteins could be assessed in the same sample of cell-extract, where as the individual cofactor requirements or stability problems of the different biosynthetic enzymes would require that separate cell-extracts be prepared for different enzyme assays. Another factor was the complexity of the LAT enzyme assay and the high instability of the LAT protein activity (188) which requires that the samples be harvested and assayed immediately for LAT activity to be detected. The LAT protein was shown by western blot analysis to be sensitive to proteolytic degradation in *S. clavuligerus* cell-extracts (data not shown) and therefore cell-extracts were immediately diluted in 2x SDS-PAGE sample buffer for storage at -20°C.

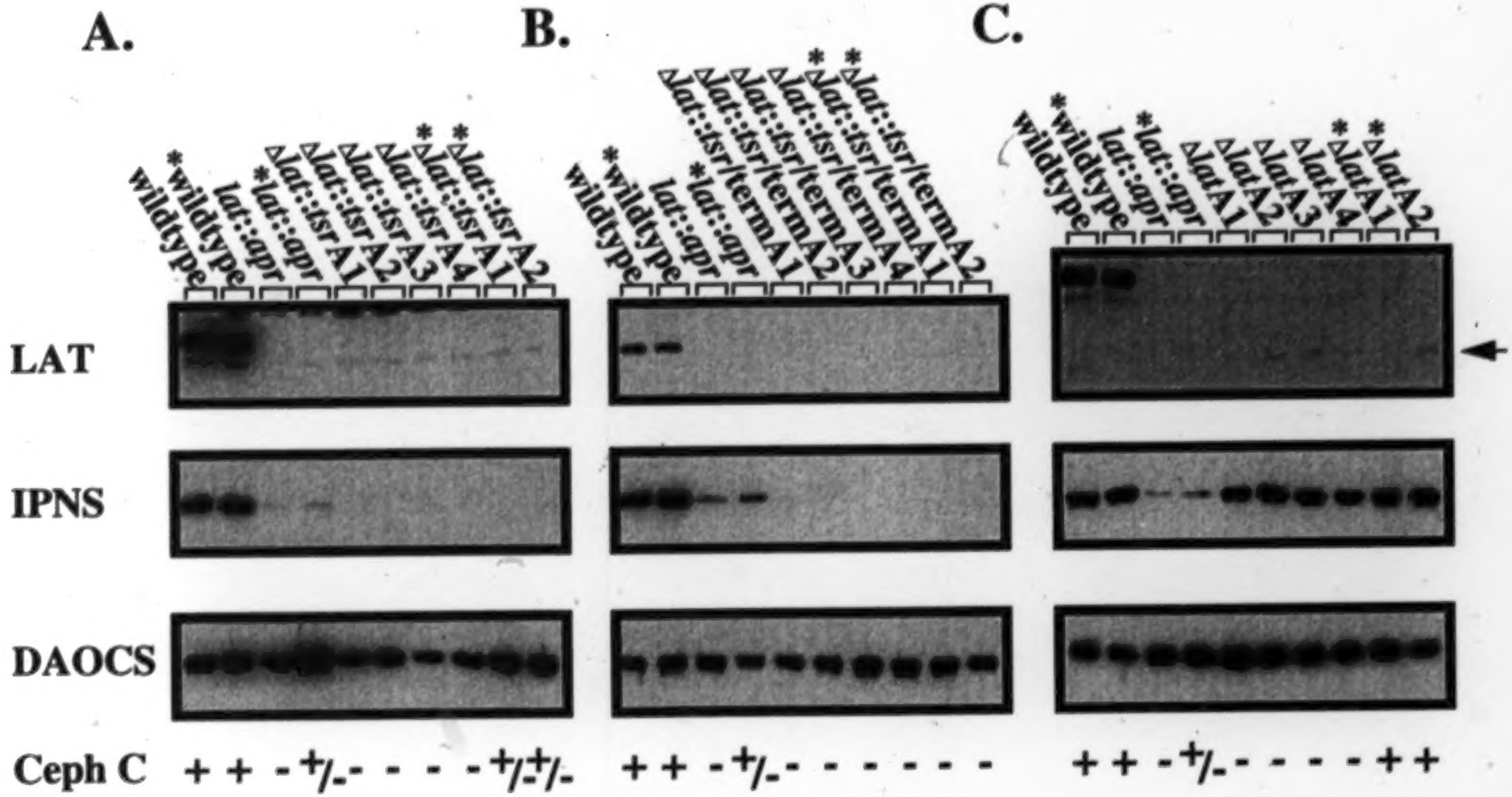
IV.5.2.1 Western blot analysis of the Δ*lat*::*tsr* mutants

The Δ*lat*::*tsr* mutants as expected did not make any LAT protein when analyzed by western analysis and compared to the wildtype *S. clavuligerus* or the *lat*::*apr* strain (Fig. 28A). These mutants all made equal amounts of DAOCS protein when compared to wildtype but did not produce detectable amounts of IPNS protein. The addition of αAA stimulated very low levels of cephamycin production from these mutants, only 1-2% of wildtype levels, but had no apparent effect on the levels of any of the biosynthetic enzymes nor did it have any effect on the proteins in the wildtype or *lat*::*apr* mutant. The fact that trace amounts of cephamycin were produced in the presence of αAA indicates that some IPNS must be produced and this could be detected during longer exposures of western blots. The low levels of IPNS protein production in the absence or presence of αAA suggests that the putative *pcbAB* promoter is not supporting the production of high levels of IPNS protein when the *lat* gene is disrupted.

IV.5.2.2 Western blot analysis of the Δ*lat*::*tsr*/term mutants

The Δ*lat*::*tsr*/term mutants also did not make any LAT protein when examined by western analysis and compared to the wildtype *S. clavuligerus* or the *lat*::*apr* strain (Fig. 28B). When compared to wildtype these mutants made equal amounts of DAOCS protein but again they did not produce any IPNS protein. Unlike the Δ*lat*::*tsr* mutants the addition of αAA did not stimulate any cephamycin production from these mutants. The addition of αAA also had no effect on the levels of any of the biosynthetic enzymes. The undetectable levels of IPNS protein production in the absence or presence of αAA suggests that the putative *pcbAB* promoter is not supporting the production of high levels of IPNS protein when the *lat* gene is disrupted.

Figure 28. Presence of the LAT, IPNS and DAOCS biosynthetic enzymes in the *lat* mutant strains. **A.** Wildtype *S. clavuligerus*, *lat*::*apr* and Δ*lat*::*tsr* partial deletion strains. **B.** Wildtype, *lat*::*apr* and Δ*lat*::*tsr*/term partial deletion strains. **C.** Wildtype, *lat*::*apr* and Δ*lat* partial deletion strains. Approximately 5 μg of cell-extract protein from each strain, harvested after 72 hours growth in TSBS or TSBS plus 2 mM αAA, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the LAT, IPNS and DAOCS proteins. The arrow designates the deleted form of the LAT protein created by an in-frame deletion of the *lat* gene. The ability of each strain to produce cephamycin (Ceph C) was determined by bioassay and the results were scored as: + for production, +/- for less than 20% of wildtype production levels and - for lack of production. The asterisks designate cultures that were supplemented with 2 mM αAA.


A.
B.
C.
*wildtype
wildtype
*lat::apr
lat::apr
Δlat::tsrA1
Δlat::tsrA2
Δlat::tsrA3
Δlat::tsrA4
*Δlat::tsrA1
*Δlat::tsrA2
Δlat::tsr/termA1
Δlat::tsr/termA2
Δlat::tsr/termA3
Δlat::tsr/termA4
*Δlat::tsr/termA1
*Δlat::tsr/termA2
ΔlatA1
ΔlatA2
ΔlatA3
ΔlatA4
*ΔlatA1
*ΔlatA2
LAT
IPNS
DAOCS
Ceph C
+ + - +/- - - - - +/- +/-
+ + - +/- - - - - - -
+ + - +/- - - - - + +

IV.5.2.3 Western blot analysis of the Δ*lat* mutants

The Δ*lat* mutants as expected did not make any *lat* protein when examined by western analysis and compared to the wildtype *S. clavuligerus* or the *lat*::*apr* strain but often after longer exposures a faint band of approximately 35 KD was detected (Fig. 28C). The internal deletion which created the Δ*lat* mutants during the repair of the *lat*::*apr* mutant was in frame and the band detected could be the deleted form of the LAT protein. When compared to the controls these mutants made equal amounts of DAOCS protein and produced amounts of IPNS protein that were equivalent to the wildtype and greatly increased over its parental strain *lat*::*apr*. The addition of αAA stimulated low levels of cephamycin production in these mutants but had no effect on the levels of any of the biosynthetic enzymes. The wildtype levels of IPNS production in the Δ*lat* mutants suggests that a functional *lat* gene is not required for the production of high levels of IPNS protein.

IV.5.3 Comparison of cephamycin biosynthetic enzyme activities in the various *lat* mutants

A single representative from each class of *lat* mutant and the wildtype and *lat*::*apr* control strains were all grown in TSBS for 48 hours. The cultures were split in half and the harvested cell pellets were sonicated in ACVS or TDE breakage buffer and assayed for certain cephamycin biosynthetic enzyme activities as well as being analyzed by western analysis and SDS-PAGE.

IV.5.3.1 Biochemical characterization of the *lat* mutants

The cell-extracts prepared in ACVS breakage buffer were assayed for ACVS enzyme activity (see II.2.4.2.1). The cell-extracts made in TDE breakage buffer were used in assays measuring IPNS, isopenicillin N epimerase (IPNE) or DAOCS activity (see II.2.4.2.2, II.2.4.2.3 and II.2.4.2.4).

The results of the assayed biosynthetic enzyme activities for the *lat* mutants are summarized in Table 13. The Δ*lat*::*tsr*A1 mutant, which was capable of producing very low levels of cephamycin when supplied with exogenous αAA, had IPNS specific activities of 14% of wildtype levels. ACVS activity was undetectable in the Δ*lat*::*tsr*A1 mutant, presumably because the low levels of activity present were difficult to detect but ACVS activity must be present as the mutants were able to make low levels of cephamycin when αAA was added. The Δ*lat*::*tsr*/termA1 mutant, which was unable to make cephamycin under any conditions, had undetectable levels of ACVS and IPNS activity. The Δ*lat*A1 mutant, which produced the highest levels of cephamycin production

Table 13. **Production of biosynthetic enzymes by the *lat* mutants**

	Biosynthetic enzyme specific activities[a]				
enzyme	wildtype	Δ*lat*::*tsr*A1	Δ*lat*::*tsr*/termA1	Δ*lat*A1	*lat*::*apr*
ACVS	100	0	0	116	17
IPNS	100	14	0	109	26
IPNE	100	72	120	110	103
DAOCS	100	53	53	88	68

[a] Specific activites values for the enzymes are given as a percentage of the wildtype specific activity, which was given a value of 100%.

after αAA feeding, had enzyme specific activities of 116% and 108% of wildtype levels for ACVS and IPNS, respectively. These levels were greatly improved over its parental strain *lat*::*apr* which only had enzyme specific activities of 17% and 26% of wildtype levels for ACVS and IPNS, respectively. All of the mutants produced near wildtype levels of IPNE but the DAOCS activities were slightly decreased.

The enzyme activity levels present in the *lat* mutants correspond to the amounts of cephamycin each strain produces in the presence of αAA.

IV.5.3.2 Western blot analysis of the *lat* mutants

As described previously, western analysis of the *lat* mutants confirmed that they were unable to produce the LAT protein (Fig. 29A). The Δ*lat*::*tsr*/termA1 mutant appeared not to produce the IPNS protein whereas the Δ*lat*::*tsr*A1mutant may make a little IPNS protein and the Δ*lat*A1 mutant made near wildtype levels of IPNS protein. This was a great improvement over its parent strain *lat::apr* which has a major reduction in the IPNS protein levels compared to wildtype and the results generated from the western analysis were very similar to the enzyme assay results. All of the mutants were shown to produce near wildtype levels of the DAOCS protein which was similar to the results of the enzyme assays for DAOCS.

Western blot analysis confirms the levels of enzyme activity detected in the various *lat* mutants.

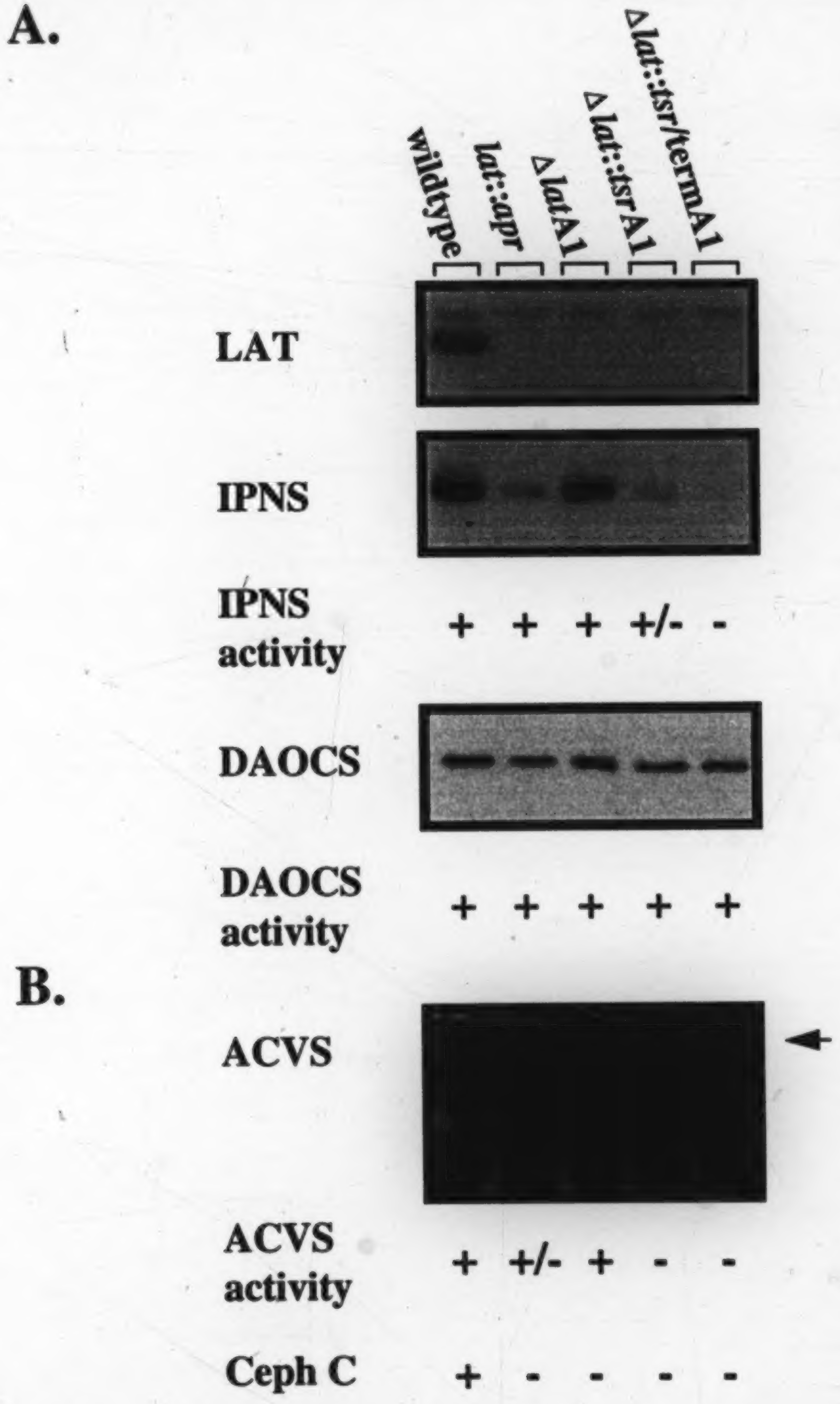
IV.5.3.3 Analysis of the *lat* mutants for production of ACVS by SDS-PAGE

The cell-extracts prepared in TDE buffer were analyzed by SDS-PAGE. The ACVS protein can be visualized by SDS-PAGE because of its very unique size, the deduced molecular weight of the *N. lactamdurans* protein was greater that 400 KD (52). If the cell-extracts are electrophoresed on a 10% gel at 50V and 4°C the ACVS protein will migrate as a sharp band very near the top of the gel. If lower percentage gels are used the band becomes much more diffuse and harder to detect.

The results generated from the SDS-PAGE analysis were very similar to the results of the enzyme assays for the mutants. Strong ACVS bands were present in the wildtype and Δ*lat*A1 lanes while reduced levels were present in the *lat::apr* mutant lane (Fig. 29B). The Δ*lat*::*tsr*A1mutant and the Δ*lat*::*tsr*/termA1 mutant appeared not to produce any ACVS protein and had undetectable amounts of ACVS activity.

SDS-PAGE analysis of the cell-extracts allows a qualitative observation of the levels of ACVS protein present in each *lat* mutant.

Figure 29. Presence of cephamycin biosynthetic enzymes in the *lat* mutant strains. **A.** Approximately 5 µg of cell-extract protein from each strain, harvested after 48 hours growth, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the LAT, IPNS and DAOCS proteins. **B.** Approximately 75 µg of cell-extract protein from each strain was separated by 10% SDS-PAGE and stained with coomassie blue. The ability of each strain to produce cephamycin (Ceph C) was determined by bioassay and the results were scored as: + for production and - for lack of production. The ability of each strain to produce the certain cephamycin biosynthetic activities was determined by enzyme assay and the results were scored as: + for production, +/- for less than 20% of wildtype production levels and - for lack of production.


A.
wildtype
lat::apr
∆*lat*A1
∆*lat::tsr*A1
∆*lat::tsr/term*A1
LAT
IPNS
IPNS activity
+ + + +/- -
DAOCS
DAOCS activity
+ + + + +
B.
ACVS
ACVS activity
+ +/- + - -
Ceph C
+ - - - -

IV.6 Complementation of *lat* mutants

It was shown that the addition of αAA to TSBS media could stimulate cephamycin production by some classes of *lat* mutants. Western blot analysis of cell-extracts from these mutants grown in the presence or absence of αAA suggested that αAA did not stimulate an increase in the level of IPNS protein for each particular class of mutant. It has been demonstrated that increasing the levels of αAA up to 4 mM continued to increase the amount of cephamycin produced by the *lat*::*apr* mutant (111), therefore it was possible that the 2 mM αAA used to supplement the TSBS may not be enough to stimulate the ACVS and IPNS protein synthesis in the Δ*lat*::*tsr* and Δ*lat*::*tsr*/term mutants.

A representative from each of the *lat* mutant sets was transformed with a pSET152-based vector (Fig. 30) containing the wildtype *lat* gene (pDA1050), the *lat* gene with its native promoter replaced by the strong constitutive *ermE**-based expression cassette (pDA1053) or the negative control plasmid containing only the *tsr* marker (pDA1000). *lat* mutants transformed with these *lat* gene constructs should make near wildtype or even greater levels of LAT protein which would result in high levels of αAA inside the cell. The pSET152-based vectors integrate at the ØC31 *attB* site within the *S. clavuligerus* genome (126) and therefore the gene(s) within the pSET152 insert would have to function *in trans*.

IV.6.1 Verification of the pSET152-based plasmid integration by genomic PCR

Since the pSET152-based plasmids used in the complementation studies integrate into the genome of *S. clavuligerus* via the *attP* site, confirmation of the presence of the plasmid or the stability of the plasmid and the insert was achieved by PCR using gDNA purified from each complemented strain as a template, together with the primers Dyl31 and Dyl32 (Fig. 31A). In this procedure, the MCS and the insert DNA cloned into the MCS of pSET152 were selectively amplified from the gDNA.

The pSET152 cloning vector was originally generated using the *E. coli* pUC19 plasmid as the backbone (27). A detailed account of the construction of pSET152 was not available but the restriction sites present in the pSET152 MCS were different from pUC19. Nonetheless the reverse and forward sequencing primers used for sequencing from either side of the MCS in pUC-like plasmids were found to be successful in sequencing the MCS pSET152 and somewhat longer oligonucleotide primers based on these sequencing primers (Dyl31 and Dyl32) were capable of specifically amplifying the MCS/insert sequence from the gDNA of the complemented mutants. If desired these oligonucleotide primers could also be used as primers for sequencing of the insert to

Figure 30. pSET152-based recombinant plasmids containing the *lat* gene. DNA fragments containing wildtype or mutant *lat* genes were cloned into the MCS of pSET152. Coding regions of the genes are indicated by closed boxes and single lines represent the intergenic regions while the dark box is the *ermE** promoter fragment.

The pDA1050 vector contains the *lat* promoter and *lat* gene while the pDA1053 vector contains the *ermE** promoter fragment attached to the *lat* gene. The pDA1000 plasmid containing the *tsr* marker was included since it serves as the negative control in these studies.

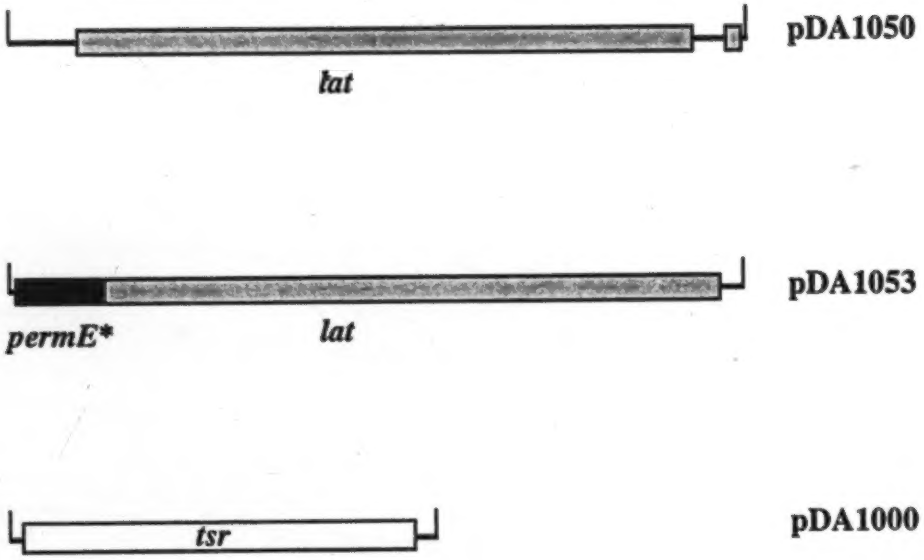
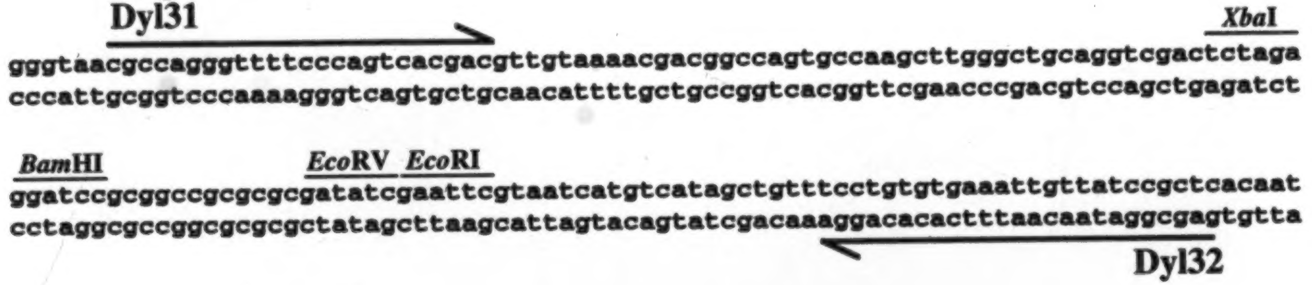

pDA1050
lat
pDA1053
permE*
lat
pDA1000
tsr

Figure 31. Confirmation of pSET152-based vector insertion by genomic PCR of transformed strains. **A.** Sequence of the pSET152 MCS. The oligonucleotide primers Dyl31 and Dyl32 allow for the specific amplification of the MCS plus the cloned insert from the pSET152-based constructs. Unique restriction sites present within the MCS are noted. **B.** Aliquots of each PCR reaction amplified from the following templates were separated on a 1.5% (w/v) agarose gel. Lane 1. pDA1000 plasmid containing the *tsr* gene purified from *E. coli*, 2. genomic DNA from pDA1000 transformed strain, 3. genomic DNA from pDA1000 transformed strain, 4. pDA1006 plasmid containing the *ccaR* gene purified from *E. coli*, 5. genomic DNA from pDA1006 transformed strain, 6. genomic DNA from pDA1006 transformed strain, 7. pDA1050 plasmid containing the *lat* gene purified from *E. coli*, 8. genomic DNA from pDA1050 transformed strain, and 9. genomic DNA from pDA1050 transformed strain.

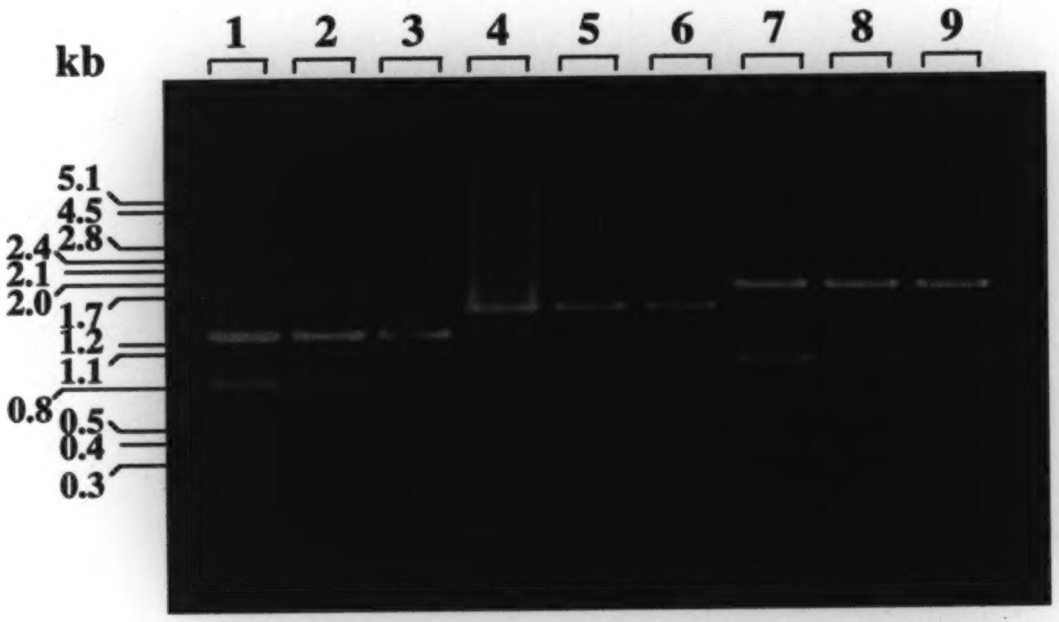
A.


Dyl31
XbaI
BamHI
EcoRV
EcoRI
Dyl32

gggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgccaagcttgggctgcaggtcgactctaga
cccattgcggtcccaaaagggtcagtgctgcaacattttgctgccggtcacggttcgaacccgacgtccagctgagatct

ggatccgcggccgcgcgcgatatcgaattcgtaatcatgtcatagctgtttcctgtgtgaaattgttatccgctcacaat
cctaggcgccggcgcgcgctatagcttaagcattagtacagtatcgacaaaggacacactttaacaataggcgagtgtta

B.


kb
1
2
3
4
5
6
7
8
9
5.1
4.5
2.8
2.4
2.1
2.0
1.7
1.2
1.1
0.8
0.5
0.4
0.3

verify that the correct insert fragment was cloned. The amplification of the MCS/insert however does not prove plasmid integration into the *att* site but successful amplification does prove integration of the pSET152-based vector into the chromosome.

Amplification of the MCS of pSET152 would generate an amplified product of only 148 bp which could be confused with a primer-dimer product. Therefore, to increase the size of the amplified product from the negative control plasmid, the *tsr* marker was cloned into pSET152 (pDA1000) which would give an amplification product of 1.3 kb. In cases where the size of the amplified product was expected to be greater than 3000 bp a combination of internal primers and MCS primers were used. This was because the Dyl31/Dyl32 primers were unsuccessful in attempts to amplify regions of this size.

Fig. 31B demonstrated that insertion of the pSET152-based vector within the chromosomal DNA could be confirmed by genomic PCR. The MCS/insert from pDA1000 (*tsr* marker), pDA1006 (*ccaR* gene) and pDA1050 (*lat* gene) were specifically amplified from the gDNA of the strains created by transformation with the pSET152-based integration vectors. The amplified PCR product from gDNA was compared to the product amplified from free plasmid isolated from *E. coli*. The product amplified should be the same size regardless of whether free plasmid or gDNA was used as the template. In all cases the size of the products amplified from the gDNA and the free plasmid were identical.

Southern blot analysis also can be used to confirm that these integrative constructs are present and stable (118) but the PCR of gDNA allows for rapid confirmation of the presence of the plasmid. If desired, the resulting PCR products could be digested with various restriction enzymes that give a diagnostic restriction pattern. However, it was felt that direct comparison of the amplified products generated from the gDNA and the plasmid isolated from *E. coli* was sufficient.

IV.6.2 Characterization of cephamycin C production by the complemented *lat* mutants

Two independent strains generated from each transformation of the *lat* mutants were grown in TSBS. Samples were taken after 48 hours and analyzed for the production of cephamycin. Also at this time the cultures were harvested and cell-extracts were prepared for western analysis.

The complementation experiments with the wildtype *lat* gene or high level *lat* expression constructs mirror the αAA feeding results, as those mutants which responded well to the αAA feeding responded well to the complementation when the *lat* gene was added (Table 14). Addition of the negative control plasmid pDA1000 to the *lat* mutants

Table 14. Production of cephamycin by complemented *lat* mutants

	cephamycin production[a]		
plasmid	Δ*lat*::*tsr*A1	Δ*lat*::*tsr*/termA1	Δ*lat*
pDA1000A	0	0	0
pDA1000B	0	0	0
mean pDA1000	0	0	0
pDA1050A	2.2	0	42.6
pDA1050B	2.3	0	51.8
mean pDA1050	2.3	0	47.2
pDA1053A	2.8	0	90.1
pDA1053B	3.9	0	104.5
mean pDA1053	3.4	0	97.3

[a] Values are calculated as a percentage of cephamycin produced by the wildtype strain which was given a value of 100%. Production values are the average of two seperate fermenations.

resulted in no complementation. The Δ*lat*::*tsr* mutants were complemented to produce an average of 2% of wildtype levels when the wildtype *lat* construct was added and up to 3.5% of wildtype levels when the high level *lat* construct was added. The Δ*lat* mutants were complemented to produce an average of 47% of wildtype levels when the wildtype *lat* construct was added and up to 97% of wildtype levels when the high level *lat* construct was added. The Δ*lat*::*tsr*/term mutants were not complemented to produce cephamycin by either construct. The addition of exogenous αAA also failed to stimulate the Δ*lat*::*tsr*/term mutant to produce cephamycin.

The levels of cephamycin produced by the complemented strains were greater than those achieved by exogenously feeding the mutant with αAA. Complementation with the *lat* gene *in trans* is expected to allow the mutants to achieve a higher intracellular αAA levels, which appear to have limited production in the αAA feeding experiments.

IV.6.3 Western blot analysis of the complemented *lat* mutants

The cell-extracts prepared from 48 hour cultures were subjected to western blot analysis using antibodies specific for the LAT, IPNS and DAOCS biosynthetic enzymes.

All sets of *lat* mutants analyzed showed the reappearance of the LAT protein when the strain was transformed with pDA1050 or pDA1053 (Fig. 32). The LAT protein was not made in strains that were transformed with the negative control plasmid pDA1000 which contains the *tsr* marker rather than the *lat* gene. All of the mutants make approximately equal amounts of DAOCS protein, with or without a functional *lat* gene being present in the cell. The Δ*lat*::*tsr*A1 mutant which was complemented to produce very low levels of cephamycin still produced undetectable amounts of the IPNS protein. The Δ*lat*::*tsr*/termA1 mutant which was not complemented to produce cephamycin also did not produce any IPNS protein. The Δ*lat*A1 mutant which was complemented to the highest levels when the *lat* constructs were added made equally high levels of the IPNS protein in the absence or presence of a functional *lat* gene.

IV.6.4 Analysis of the complemented *lat* mutants for production of ACVS by SDS-PAGE

The cell-extracts prepared from 48 hour cultures were analyzed by SDS-PAGE under conditions which allow detection of the ACVS protein.

The addition of a functional *lat* gene did not stimulate the production of ACVS protein by the Δ*lat*::*tsr*A1 and Δ*lat*::*tsr*/termA1 mutants (Fig. 33). Only the Δ*lat*A1 mutant made detectable amounts of the ACVS protein and the levels of ACVS protein do not appear to change in the presence of a functional *lat* gene.

Figure 32. Western analysis of *lat* mutants complemented with pSET152-based recombinant plasmids containing the *lat* gene. Western blot analysis and bioassay for production of cephamycin were used to assess complementation. The designation A or B represent two independent strains generated during the same transformation. Approximately 5 μg of cell-extract protein from each strain, harvested after 48 hours growth, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the LAT, IPNS and DAOCS proteins. The arrow designates a deleted form of the LAT protein. The ability of each strain to produce cephamycin (Ceph C) was determined by bioassay and the results were scored as: + for production, +/- for less than 20% of wildtype production levels and - for lack of production.

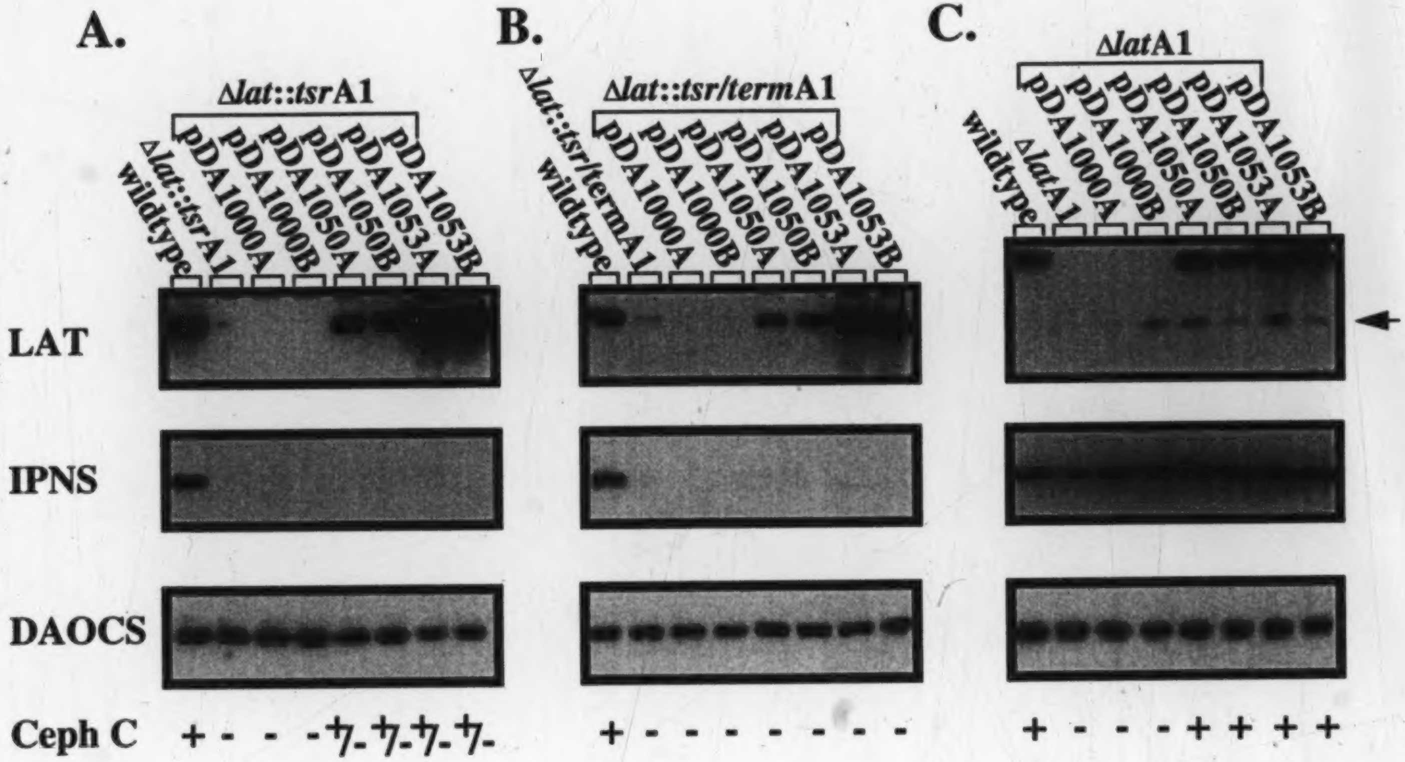
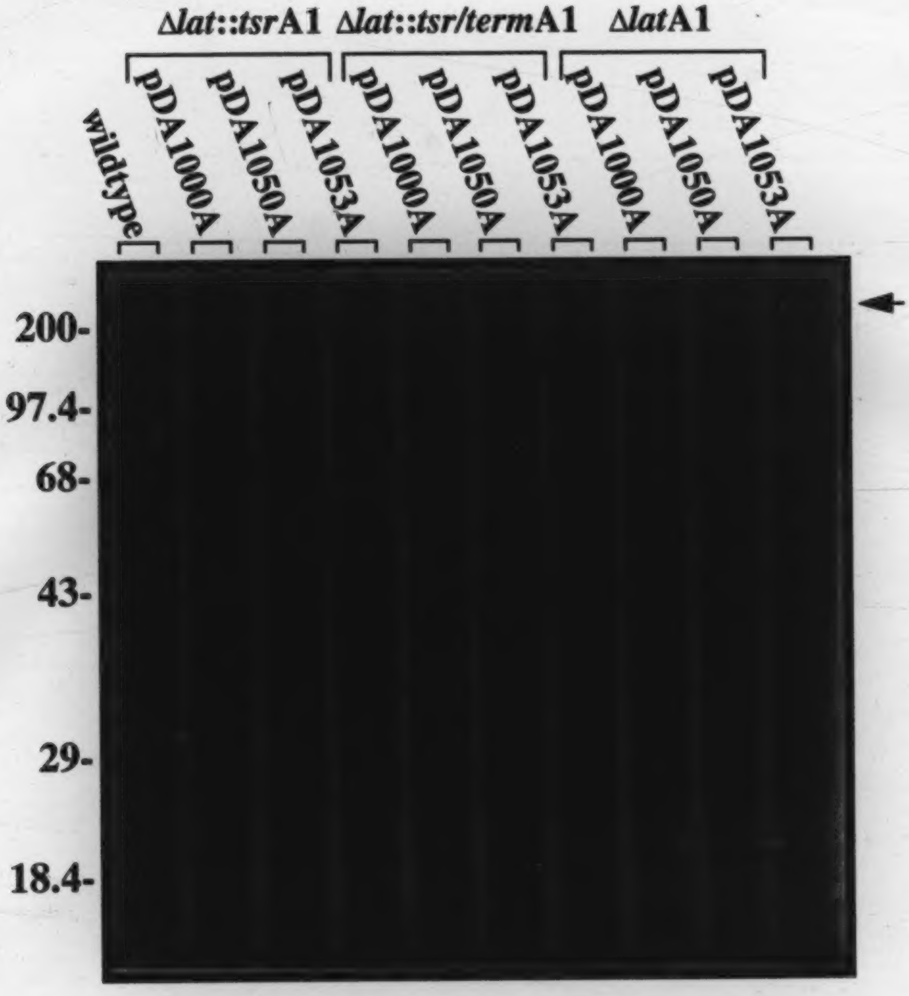

A.
Δlat::tsrA1
wildtype
Δlat::tsrA1
pDA1000A
pDA1000B
pDA1050A
pDA1050B
pDA1053A
pDA1053B
B.
Δlat::tsr/termA1
wildtype
Δlat::tsr/termA1
pDA1000A
pDA1000B
pDA1050A
pDA1050B
pDA1053A
pDA1053B
C.
ΔlatA1
wildtype
ΔlatA1
pDA1000A
pDA1000B
pDA1050A
pDA1050B
pDA1053A
pDA1053B
LAT
IPNS
DAOCS
Ceph C
+ - - - +/- +/- +/- +/-
+ - - - - - - -
+ - - - + + + +

Figure 33. SDS-PAGE analysis of *lat* mutants complemented with pSET152-based recombinant plasmids containing the *lat* gene. Approximately 75 µg of cell-extract protein from each strain, harvested after 48 hours growth, was separated by 10% SDS-PAGE and stained with coomassie blue. Molecular weight markers were run in adjacent lanes and their molecular weight sizes in kilodaltons were labeled beside the gel in this figure.


Δlat::tsrA1
Δlat::tsr/termA1
ΔlatA1
wildtype
pDA1000A
pDA1050A
pDA1053A
pDA1000A
pDA1050A
pDA1053A
pDA1000A
pDA1050A
pDA1053A
200-
97.4-
68-
43-
29-
18.4-

IV.7 Promoter probe analysis of *lat* mutants

The *lat* mutants created by homologous recombination do not make any LAT protein as judged by western blot analysis, except for the Δ*lat* mutant which may make low levels of a LAT protein containing an internal deletion. It was of interest to determine whether the *lat* promoter remained active in these αAA-free cells or whether αAA played a role in activating or stimulating the transcription of the *lat* gene.

A representative from each of the *lat* mutant sets created in this study was transformed with the promoter probe plasmid pDA2001 (see IV.3.1). This construct contains only the *lat* promoter fragment upstream of a promoterless *xylE* gene and therefore would be unable to complement the *lat* mutant strains.

IV.7.1 Analysis of *lat* promoter activity in *lat* mutants

Two independent strains generated from each transformation were grown for 48 hours in TSBS. The harvested cell pellets were resuspended in phosphate breakage buffer, disrupted by sonication and the cell-extracts were then assayed for C23O activity.

The C23O specific activity in these mutant backgrounds was measured and compared to the strength of the *lat* promoter in wildtype *S. clavuligerus* (Table 15). All of the *lat* mutant sets have near wildtype levels of C23O specific activity or greater, demonstrating that the *lat* promoter was definitely functional in these *lat* mutant strains.

Transcription from the *lat* promoter was not dependent upon a functional *lat* gene allowing the production of αAA, suggesting that αAA is not autostimulatory for transcription of the *lat* gene.

Table 15. Promoter probe analysis of *lat* promoter in *lat* mutants

	Catechol dioxygenase specific activity[a]			
strain	strain A[b]	strain B[b]	average	average / wildtype [c]
wildtype	50	70	60	1
Δ*lat*::*tsr*A1	53	49	51	85
Δ*lat*::*tsr*/termA1	60	41	51	85
Δ*lat*A1	98	87	92	153

[a] C23O specific activity values are given as EU x 10^{-6}/mg protein.

[b] Specific activities are an average of two seperate fermentations for each strain.

[c] Values are calculated as a percentage of the specific activity obtained for the wildtype *S. clavuligerus* transformed with the pDA2001 plasmid.

V. Results. Regulation of β-lactam biosynthesis by the CcaR protein

Database searches for homologs of the CcaR protein detected many transcriptional activators from *Streptomyces* involved in the regulation of antibiotic biosynthetic pathways. The deduced amino acid sequences from these proteins and the *Mycobacteria* homologs were quite similar to the CcaR amino acid sequence (see Fig. 13). Walters et al. (221) characterized mutants unable to produce both clavulanic acid and cephamycin C and demonstrated that complementation of the mutant strains with the *dclX* (*ccaR*) gene rescued production of both antibiotics. Later Perez-Llarena et al. (174) showed that mutation of the *ccaR* gene eliminated the biosynthesis of both cephamycin and clavulanic acid. I have also investigated this regulatory gene in order to better understand its mechanism of regulation and the role of the two uncharacterized open reading frames (ORFs) *orf11* and *blp* which lie downstream of *ccaR*, in the regulation of β-lactam biosynthesis in *S. clavuligerus*.

V.1 Mutation of the *ccaR* gene or neighboring genes by gene replacement

In order to study the regulation of β-lactam biosynthesis by the CcaR protein it was necessary to create mutants of *S. clavuligerus* which lacked the CcaR protein. The Dcl mutants isolated by Walters et al. (221) were derivatives of high titre industrial production strains and thus were not available. As well, these mutant strains were generated during the highly mutagenic strain improvement protocols and likely contain multiple mutations which may have confused the results of later experiments. Therefore, it was desired to create *ccaR* mutant strains that contain a disrupted or deleted *ccaR* gene but were otherwise isogenic to wildtype *S. clavuligerus*.

The presence of the *orf11* and *blp* genes in the cephamycin cluster was not expected and an obvious role for these genes in antibiotic biosynthesis was not known. A possible role for the Blp protein could not be deduced since database searches showed homology to only BLIP, a proteinaceous β-lactamase inhibitor from *S. clavuligerus* and the *blp* gene does not contain a homologous gene in any of the other β-lactam producing strains (see III.8.1). The role of the Orf11 protein was also uncertain since the protein had no homology to any known proteins in the database searches. However, the Orf11 protein may be involved in the production of clavulanic acid and/or clavams since the *orf11* gene appears to have a homologous gene in *S. jumonjinensis* and *S. lipmanii* (see III.8.1), which produce clavulanic acid or clavulanic acid precursors, respectively (19, 69). If the *pcd* gene does code for the piperideine-6-carboxylate dehydrogenase [PCD, (55)] then all of the genes required to produce the biosynthetic enzyme activities necessary for cephamycin biosynthesis have been accounted for in the *S. clavuligerus* cephamycin cluster. Therefore, the *orf11* and *blp* genes are more likely to play a role in the regulation of antibiotic production or resistance to the antibiotics produced.

All of the gene targeting constructs described below were prepared as parallel sets in which the antibiotic resistance marker was inserted in both orientations relative to the target gene(s) being disrupted or deleted. This makes it possible to determine if the mutation to the target gene(s) was responsible for the altered phenotype or if the antibiotic resistance marker had caused a polar mutation affecting other genes within an operon. This precaution regarding polar mutations may have been unnecessary since the *ccaR* and *blp* genes were expressed as monocistronic messages and a transcript for *orf11* was undetectable in trypticase soy broth (174).

The mutations were generated by a procedure inviting homologous recombination between the targeting vector carrying the mutagenic gene and the parental chromosome (see II.2.5). The targeting vector crosses over into the chromosome via the homologous DNA insert and then crosses back out to complete the double crossover. When the single crossovers occur on either side of the mutated gene within the targeting vector the mutant gene would replace the wildtype gene in the chromosome. Putative mutants were selected for maintenance of the apramycin resistant phenotype (apr^R; disruption marker) or the thiostrepton resistant phenotype (tsr^R; deletion marker) as well as the loss of the hygromycin resistant phenotype (hyg^R; shuttle vector marker). Two mutants for each gene replacement experiment were chosen for both orientations of the antibiotic resistance marker used.

The mutations were confirmed by Southern blot analysis of restriction enzyme digested genomic DNA (gDNA) purified from parental and mutant strains. For simplification two mutants from each mutant background (one from each orientation) were included on one final Southern blot. Southern analysis characterizing all four individual mutants created in a particular set of gene targeting experiments was done previously and all mutants from a particular set had the identical hybridization pattern.

V.1.1 Construction of the *apr* cassette

The apramycin resistance (*apr*) marker was removed from pOW309 and inserted into pUC119 to create pDA515. The *apr* marker functions as a bifunctional resistance marker, providing both *E. coli* and *S. clavuligerus* resistance to apramycin (apr). The plasmid was linearized at the unique *Sph*I site within the multiple cloning site (MCS) of pDA515 and the stop codon adapter oligonucleotide Dyl19 was inserted immediately upstream of the *apr* marker to create pDA516. The oligonucelotide contained an *Spe*I recognition site and stop codons in all three reading frames. The *apr* marker and adapter oligonucleotide (*apr* cassette) were removed from pDA516 and cloned into pBluescript KS(+) to give pDA517.

The *apr* cassette was removed from pDA517 by *Hind*III and *Sma*I digestion and the fragment ends were made blunt by treatment with Klenow polymerase (Fig. 34). Insertion of the *apr* cassette was selected by plating the *E. coli* transformation mixture on plates containing apr and selecting for colonies with an apr^R phenotype. Orientation of the *apr* cassette inserted into the target gene was determined by digestion with *Eco*RI and an enzyme recognizing a unique site within the MCS of the plasmid. When the *apr* cassette was to be used to terminate translation, the *apr* marker was removed as a *Pst*I fragment and the remaining plasmid was self-ligated to leave behind the 37 bp translational stop fragment (TSF) containing amber stop codons in all three reading frames within the target gene. If the TSF was placed in frame with the target gene seven extra amino acids would be added to the target gene before the stop codon was reached. If the TSF was out of frame by one base pair or two base pairs six or four extra amino acids would be added to the target gene, respectively. It was felt that the small size of the TSF would not negatively effect the transcription of downstream genes within a polycistronic transcript.

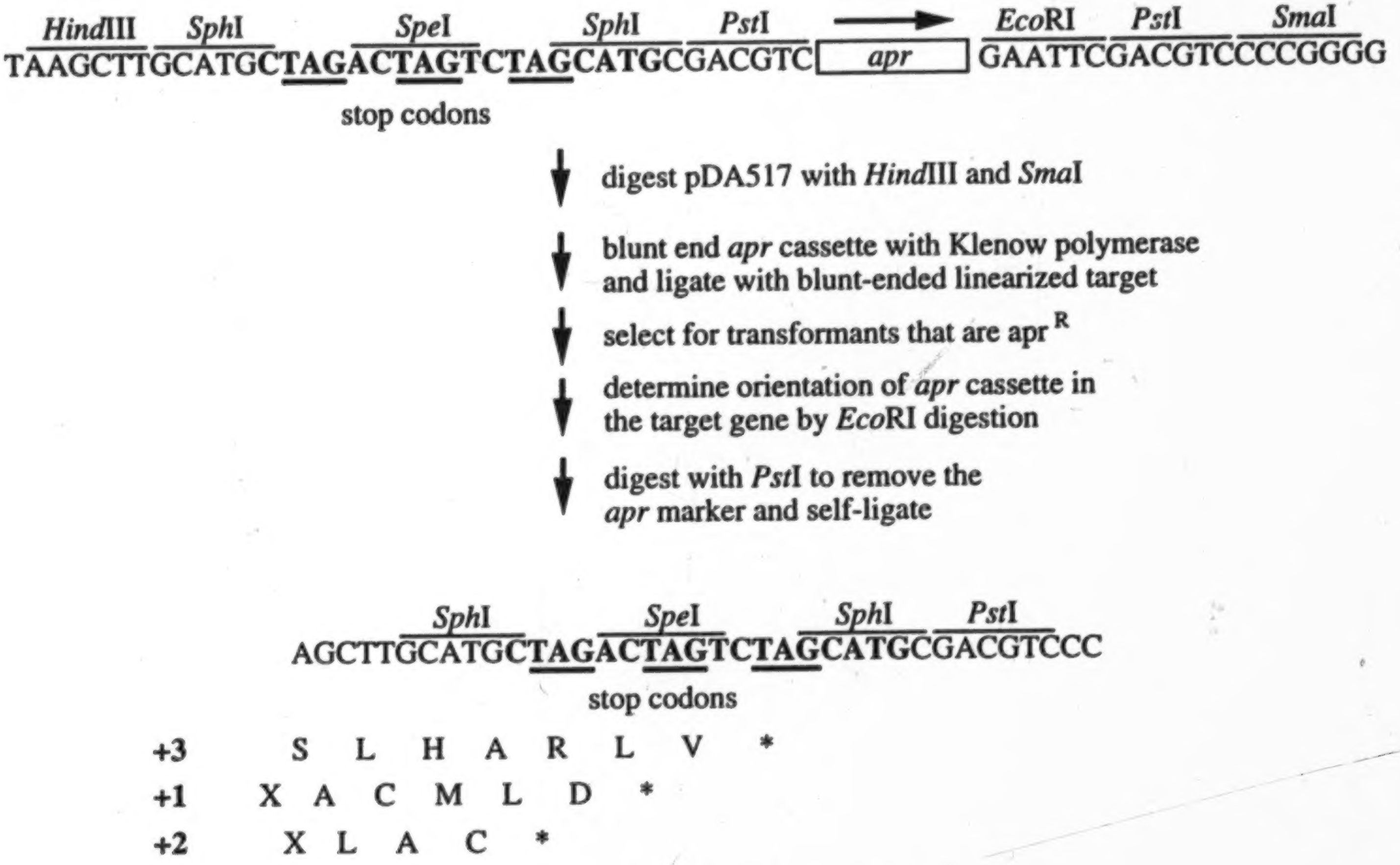
V.1.2 Disruption of the *cob* genes with the *apr* cassette

The primary goal of the research described in this section was to study regulation of antibiotic biosynthesis by the CcaR transcriptional activator but mutants of the *orf11* and *blp* genes were also desired. Therefore three different sets of targeting vectors were prepared. This group of genes (*c̲caR*, *o̲rf11* and *b̲lp* genes) are referred to as the *cob* genes for convenience. The first set of disruption mutants resulted from insertion of the *apr* cassette into the unique *Eco*ICRI site within the *ccaR* coding region. The second set of disruption mutants resulted from the insertion of the *apr* cassette into the unique *Nco*I site within the *orf11* coding region. The final set of disruption mutants was to result from the insertion of the *apr* cassette into the unique *Not*I site within the *blp* coding region. However, these mutants could not be obtained in wildtype *S. clavuligerus*.

V.1.2.1 *ccaR* disruption with the *apr* cassette (*ccaR::apr*)

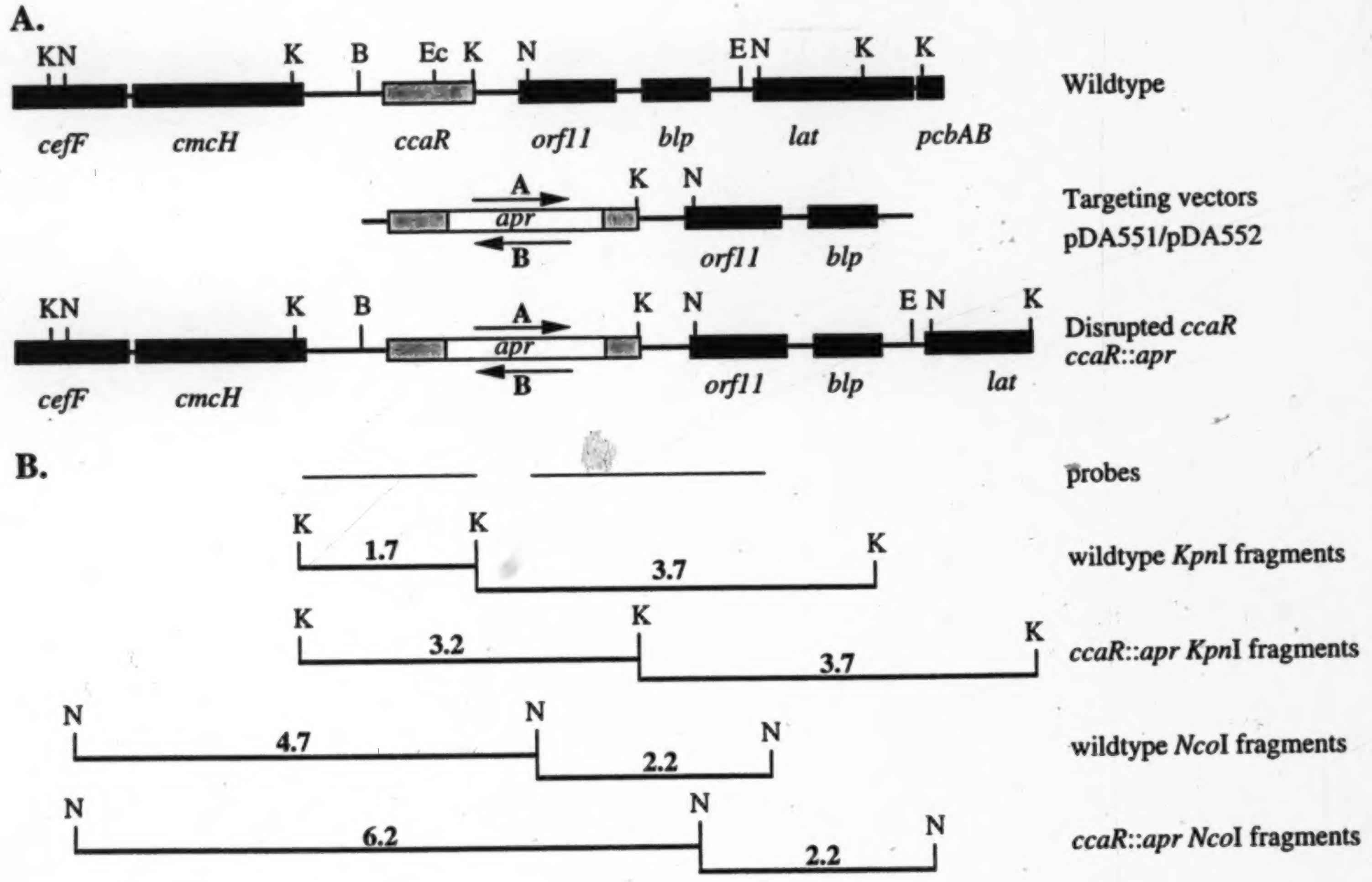
The first set of disruption mutants resulted from the insertion of the *apr* cassette into the unique *Eco*ICRI site of the *ccaR* gene to create a *ccaR* disruption mutant (Fig. 35). The *apr* cassette was inserted in both orientations to determine if a polar mutation

Figure 34. Diagrammatic representation of the *apr* cassette containing the bifunctional *apr* marker. After insertion of the blunt-ended *apr* cassette into the targeted gene this plasmid can be made into a shuttle vector to be used for a gene disruption by homologous recombination. Alternatively, the *apr* marker can be removed from the plasmid by *Pst*I digestion and self-ligation to create a mutant gene as the residual 37 bp translation stop fragment will generate stop codons in all three reading frames. The number of amino acids added to the mutant protein before the stop codon was reached can be estimated by determining in what frame the translation stop fragment was inserted relative to the target gene. The sequence of the Dyl19 oligonucleotide is present in bold text while the amber stop codons present in the oligonucleotide are underlined. Restriction sites present within the *apr* cassette are noted.


HindIII
SphI
SpeI
SphI
PstI
EcoRI
PstI
SmaI
TAAGCTTGCATGCTAGACTAGTCTAGCATGCGACGTC
apr
GAATTCGACGTCCCCGGGG
stop codons
digest pDA517 with HindIII and SmaI
blunt end apr cassette with Klenow polymerase and ligate with blunt-ended linearized target
select for transformants that are apr R
determine orientation of apr cassette in the target gene by EcoRI digestion
digest with PstI to remove the apr marker and self-ligate
SphI
SpeI
SphI
PstI
AGCTTGCATGCTAGACTAGTCTAGCATGCGACGTCCC
stop codons
+3 S L H A R L V *
+1 X A C M L D *
+2 X L A C *

37 bp translation stop fragment generating stop codons in all three reading frames

Figure 35. Strategy for insertional inactivation of the *ccaR* gene and replacement with the *apr* cassette. **A.** The dark shaded boxes represent ORFs adjacent to *ccaR* in the chromosome and the lighter shaded boxes represent the *ccaR* gene within the wildtype cephamycin cluster. The targeting vectors pDA551 and pDA552 contained the *cob* gene DNA fragment in which the *ccaR* coding region was interrupted with the *apr* cassette, represented by the clear box, inserted at the *Eco*ICRI site. The *apr* marker is shown to be present in both orientations (A or B). The mutant chromosome results as the disrupted *ccaR* gene containing the *apr* cassette replaces the wildtype gene in the chromosome. **B.** Southern hybridization pattern of *Kpn*I or *Nco*I digested genomic DNA from wildtype and *ccaR*::*apr* mutants using the 1.7 kb *Kpn*I fragment and the 2.2 kb *Nco*I fragment as probes. Abbreviations for restrictions sites: B for *Bam*HI, E for *Eco*RI, Ec for *Eco*ICRI, K for *Kpn*I and N for *Nco*I.


A.
KN
K
B
Ec
K
N
EN
K
K
cefF
cmcH
ccaR
orf11
blp
lat
pcbAB
Wildtype
A
apr
B
K
N
orf11
blp
Targeting vectors
pDA551/pDA552
KN
K
B
A
apr
B
K
N
E N
K
cefF
cmcH
orf11
blp
lat
Disrupted ccaR
ccaR::apr
B.
probes
K
1.7
K
3.7
K
wildtype KpnI fragments
K
3.2
K
3.7
K
ccaR::apr KpnI fragments
N
4.7
N
2.2
N
wildtype NcoI fragments
N
6.2
N
2.2
N
ccaR::apr NcoI fragments

was generated which affected nearby genes. The plasmid constructs pDA551 and pDA552 carried a stretch of *S. clavuligerus* DNA containing the *cob* genes in which the *ccaR* coding region was interrupted with the *apr* cassette. Each targeting vector was introduced into protoplasts of wildtype *S. clavuligerus* by transformation and two transformants were allowed to sporulate in the absence of antibiotic selection to promote the loss of free plasmid. Putative mutants were selected for the maintenance of the apr^R phenotype but loss of the hyg^R phenotype.

The mutations were confirmed by Southern blot analysis. The digested gDNA was hybridized with the 1.7 kb *Kpn*I fragment containing the end of the *cmcH* gene and almost the entire *ccaR* gene and the 2.2 kb *Nco*I fragment containing the *orf11* and *blp* genes and the *lat* promoter (Fig. 35). The wildtype *Kpn*I digested gDNA gave hybridizing bands at 3.7 and 1.7 kb (Fig. 39, lane 1) while the *ccaR*::*apr* mutants gave 3.7 and 3.2 kb hybridizing bands (lanes 2-3). The wildtype *Nco*I digested gDNA gave hybridizing bands of 4.7 and 2.2 kb (lane 11), while the *ccaR*::*apr* mutants gave hybridizing bands of 6.2 kb and 2.2 kb (lanes 12-13).

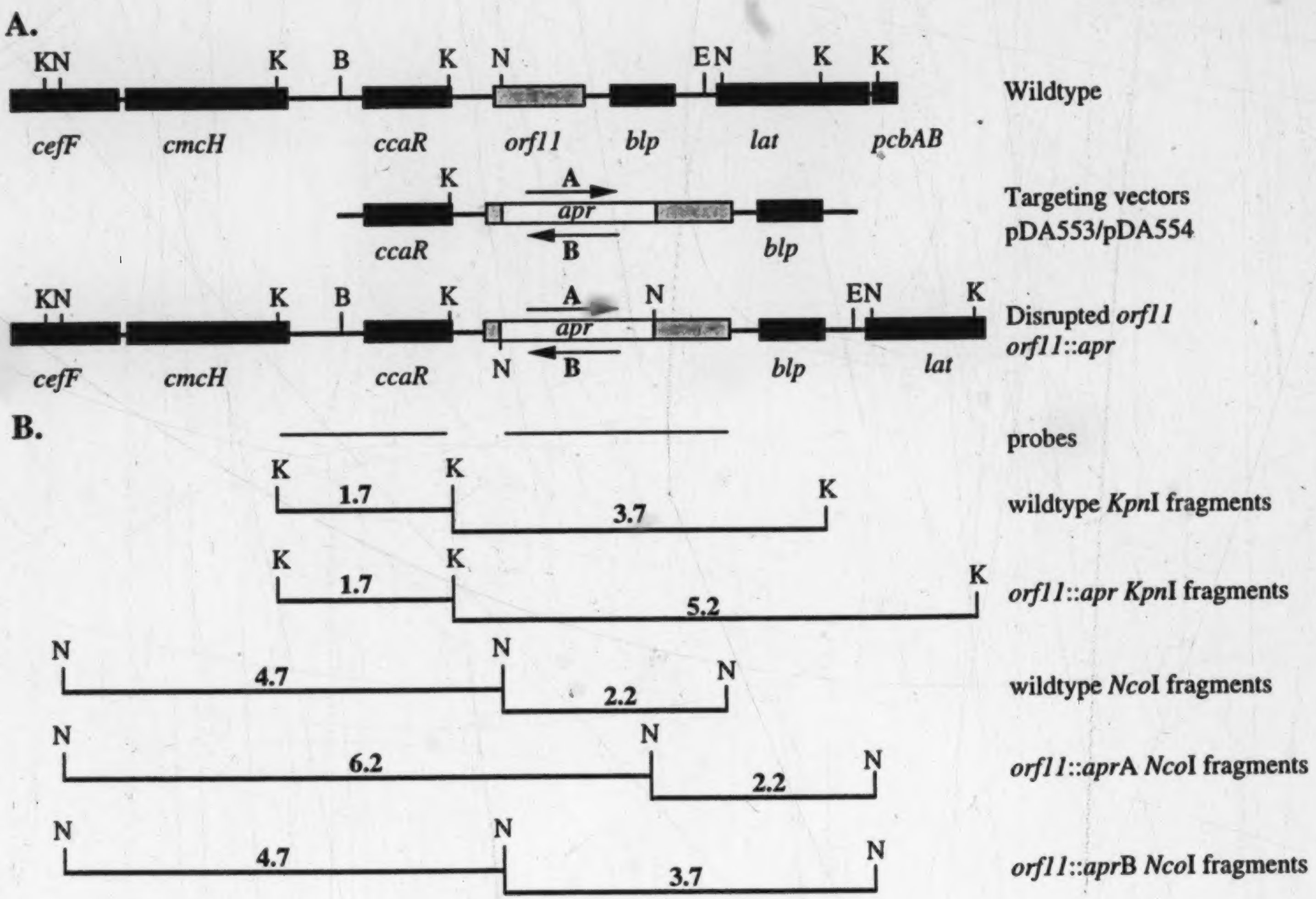
The above results were all consistent with a double crossover event having occurred in the *ccaR*::*apr*A and *ccaR*::*apr*B mutants of *S. clavuligerus*.

V.1.2.2 *orf11* disruption with the *apr* cassette (*orf11*::*apr*)

The second set of disruption mutants constructed inside the *cob* genes was the insertion of the *apr* cassette into the unique *Nco*I site of *orf11* to create an *orf11* disruption mutant (Fig. 36). Again the *apr* cassette was inserted in both orientations to avoid any polar effects on nearby genes. The targeting vectors pDA553 or pDA554, carrying a stretch of *S. clavuligerus* DNA containing the *cob* genes in which the *orf11* coding region was interrupted with the *apr* cassette, were transformed into wildtype *S. clavuligerus* protoplasts. Two transformants were allowed to sporulate in the absence of antibiotic selection to promote the loss of free plasmid. Putative mutants were selected for maintenance of the apr^R phenotype but loss of the hyg^R phenotype.

The mutations were confirmed by Southern blot analysis. The digested gDNA was hybridized with the 1.7 kb *Kpn*I fragment and the 2.2 kb *Nco*I fragment (Fig. 36). The wildtype *Kpn*I digested gDNA gave hybridizing bands of 3.7 and 1.7 kb (Fig. 39, lane 1) while the *orf11*::*apr* mutants gave 5.2 and 1.7 kb hybridizing bands (lanes 4-5). The wildtype *Nco*I digested gDNA gave hybridizing bands of 4.7 and 2.2 kb (lane 11), while the *orf11*::*apr*A1 mutant gave hybridizing bands of 6.2 and 2.2 kb (lane 14) and the while the *orf11*::*apr*B1 mutant gave hybridizing bands of 4.7 and 3.7 kb (lane 15). The *orf11*::*apr* mutants differ when they were digested with *Nco*I because the *Nco*I site was

Figure 36. Strategy for insertional inactivation of the *orf11* gene and replacement with the *apr* cassette. **A.** The dark shaded boxes represent ORFs adjacent to *orf11* in the chromosome and the lighter shaded boxes represent the *orf11* gene within the wildtype cephamycin cluster. The targeting vectors pDA553 and pDA554 contained the *cob* gene DNA fragment in which the *orf11* coding region was interrupted with the *apr* cassette, represented by the clear box, inserted at the *Nco*I site. The *apr* marker is shown to be present in both orientations (A or B). The mutant chromosome results as the disrupted *orf11* gene containing the *apr* cassette replaces the wildtype gene in the chromosome. **B.** Southern hybridization pattern of *Kpn*I or *Nco*I digested genomic DNA from wildtype and *orf11*::*apr* mutants using the 1.7 kb *Kpn*I fragment and the 2.2 kb *Nco*I fragment as probes. Abbreviations for restrictions sites: B for *Bam*HI, E for *Eco*RI, K for *Kpn*I and N for *Nco*I.


A.
KN
K
B
K
N
EN
K
K
Wildtype
cefF
cmcH
ccaR
orf11
blp
lat
pcbAB
K
A
apr
B
Targeting vectors
pDA553/pDA554
ccaR
blp
KN
K
B
K
A
N
EN
K
apr
N
B
Disrupted orf11
orf11::apr
cefF
cmcH
ccaR
blp
lat
B.
probes
K
K
K
1.7
3.7
wildtype KpnI fragments
K
K
K
1.7
5.2
orf11::apr KpnI fragments
N
N
N
4.7
2.2
wildtype NcoI fragments
N
N
N
6.2
2.2
orf11::aprA NcoI fragments
N
N
N
4.7
3.7
orf11::aprB NcoI fragments

not destroyed when the blunt-ended *apr* cassette was inserted. The *Nco*I site was regenerated on one end of the *apr* cassette and which side the *Nco*I site was regenerated would vary depending on which orientation the *apr* cassette inserted.

The above results were all consistent with a double crossover event having occurred in the *orf11*::*apr*A and *orf11*::*apr*B mutants of *S. clavuligerus*.

V.1.2.3 *blp* disruption with the *apr* cassette (*blp*::*apr*)

The final set of disruption mutants constructed inside the *cob* genes involved the insertion of the *apr* cassette into the unique *Not*I site of *blp* to create a *blp* disruption mutant. Again the *apr* cassette was inserted in both orientations to avoid any polar effects on nearby genes. The plasmid constructs pDA555 and pDA556 carried a stretch of *S. clavuligerus* DNA containing the *cob* genes in which the *blp* coding region was interrupted with the *apr* cassette. The targeting vectors were transformed into wildtype *S. clavuligerus* protoplasts and two transformants were allowed to sporulate in the absence of antibiotic selection to promote the loss of free plasmid. Screening for putative mutants that had maintained the apr^R phenotype but lost the hyg^R phenotype was unsuccessful. Rather the colonies had phenotypes suggesting that they had either lost the plasmid without any recombination occurring (apr^S/hyg^S) or retained the plasmid (apr^R/hyg^R).

Similar phenotypes were observed during attempts to mutate the *pcbR* gene in wildtype *S. clavuligerus* (169). Mutants of *pcbR* eventually were made in the cephamycin non-producing strain NTG 19 (166). PcbR was apparently essential for survival in the presence of the high levels of cephamycin produced by the wildtype strain. Since a physiological function for Blp or the homologous BLIP are unknown it was thought that Blp may be involved in resistance and this would prevent *blp* mutants from being made in wildtype *S. clavuligerus*. Attempts were made to create a *blp* mutant in the antibiotic non-producing Δ*ccaR*::*tsr*A1 strain (see below). Both orientations of the *blp*::*apr* mutants were quickly created in this antibiotic non-producing background but these mutants obviously could not be further characterized for the *blp* mutation's effects on antibiotic production. Further attempts to create *blp* mutants in wildtype *S. clavuligerus* were not made since *blp* mutants were constructed in an antibiotic producing background through other means (see V.8.2).

V.1.3 Deletion of *cob* genes and replacement with the *tsr* marker

Deletion mutants were also created for characterization of the roles of the *cob* genes in the regulation of antibiotic production. The deletion mutants were created with the *tsr* marker so that the mutant strains could be complemented with recombinant plasmids created using the pSET152 integrating vector. Two different sets of tsr^R deletion mutants were created. The first set of deletion mutants resulted from the deletion of the *Bam*HI/*Nru*I fragment containing the *ccaR* gene. The final set of deletion mutants resulted from the deletion of the *Bam*HI/*Eco*RI fragment containing the *cob* genes.

V.1.3.1 *ccaR* deletion and replacement with the *tsr* marker (Δ*ccaR*::*tsr*)

The first set of two deletion mutant strains created has a deletion of the *Bam*HI/*Nru*I fragment containing the entire *ccaR* gene, which was replaced with the *tsr* marker to create a *ccaR* deletion mutant (Fig. 37). The *tsr* marker was inserted in both orientations to avoid any polar effects on nearby genes. The plasmid constructs pDA557 and pDA558 carried a contiguous stretch of *S. clavuligerus* DNA which extends from *cefF* to *blp* except that the 1.4 kb *Bam*HI/*Nru*I fragment containing the *ccaR* gene was removed and replaced with the *tsr* marker. The targeting vectors were transformed into wildtype *S. clavuligerus* protoplasts and two transformants were allowed to sporulate in the absence of antibiotic selection to promote the loss of free plasmid. Putative mutants were selected for maintenance of the tsr^R phenotype but loss of the hyg^R phenotype.

The mutations were confirmed by Southern blot analysis. The digested gDNA was hybridized with the 1.7 kb *Kpn*I fragment and the 2.2 kb *Nco*I fragment (Fig. 37). The wildtype *Kpn*I digested gDNA gave hybridizing bands at 3.7 and 1.7 kb (Fig. 39, lane 1) while the Δ*ccaR*::*tsr* mutants gave a 5.0 kb hybridizing band (lanes 6-7). The wildtype *Nco*I digested gDNA gave hybridizing bands of 4.7 and 2.2 kb (lane 11), while the Δ*ccaR*::*tsr* mutants gave hybridizing bands of 4.4 and 2.2 kb (lanes 16-17).

The above results were all consistent with a double crossover event having occurred in the Δ*ccaR*::*tsr*A and Δ*ccaR*::*tsr*B mutants of *S. clavuligerus*.

V.1.3.2 *cob* deletion and replacement with the *tsr* marker (Δ*cob*::*tsr*)

The second set of two deletion mutant strains constructed has a deletion of the *Bam*HI/*Eco*RI fragment containing the *cob* genes, which was replaced with the *tsr* marker to create a *cob* deletion mutant (Fig. 38). Again the *tsr* marker was inserted in both orientations to avoid any polar effects on nearby genes. The plasmid constructs pDA559 and pDA560 carried a contiguous stretch of *S. clavuligerus* DNA which extends from *cefF* to *pcbAB* except that the 3.7 kb *Bam*HI/*Eco*RI fragment containing the *cob* genes was removed and replaced with the *tsr* marker. The targeting vectors were transformed into wildtype *S. clavuligerus* protoplasts and two transformants were allowed to sporulate

Figure 37. Strategy for the deletion of the *ccaR* gene and replacement with the *tsr* marker. **A.** The dark shaded boxes represent ORFs adjacent to *ccaR* in the chromosome and the lighter shaded boxes represent the *ccaR* gene within the wildtype cephamycin cluster. The targeting vectors pDA557 and pDA558 contained a contiguous stretch of *S. clavuligerus* DNA which extends from *cefF* to *blp* except that the 1.4 kb *Bam*HI/*Nru*I fragment containing the *ccaR* gene was removed and replaced with the *tsr* marker, represented by the clear box. The *tsr* marker is shown to be present in both orientations (A or B). The mutant chromosome results as the *tsr* marker replaces the wildtype gene in the chromosome. **B.** Southern hybridization pattern of *Kpn*I or *Nco*I digested genomic DNA from wildtype and Δ*ccaR*::*tsr* mutants using the 1.7 kb *Kpn*I fragment and the 2.2 kb *Nco*I fragment as probes. Abbreviations for restrictions sites: B for *Bam*HI, E for *Eco*RI, K for *Kpn*I, N for *Nco*I and Nr for *Nru*I.

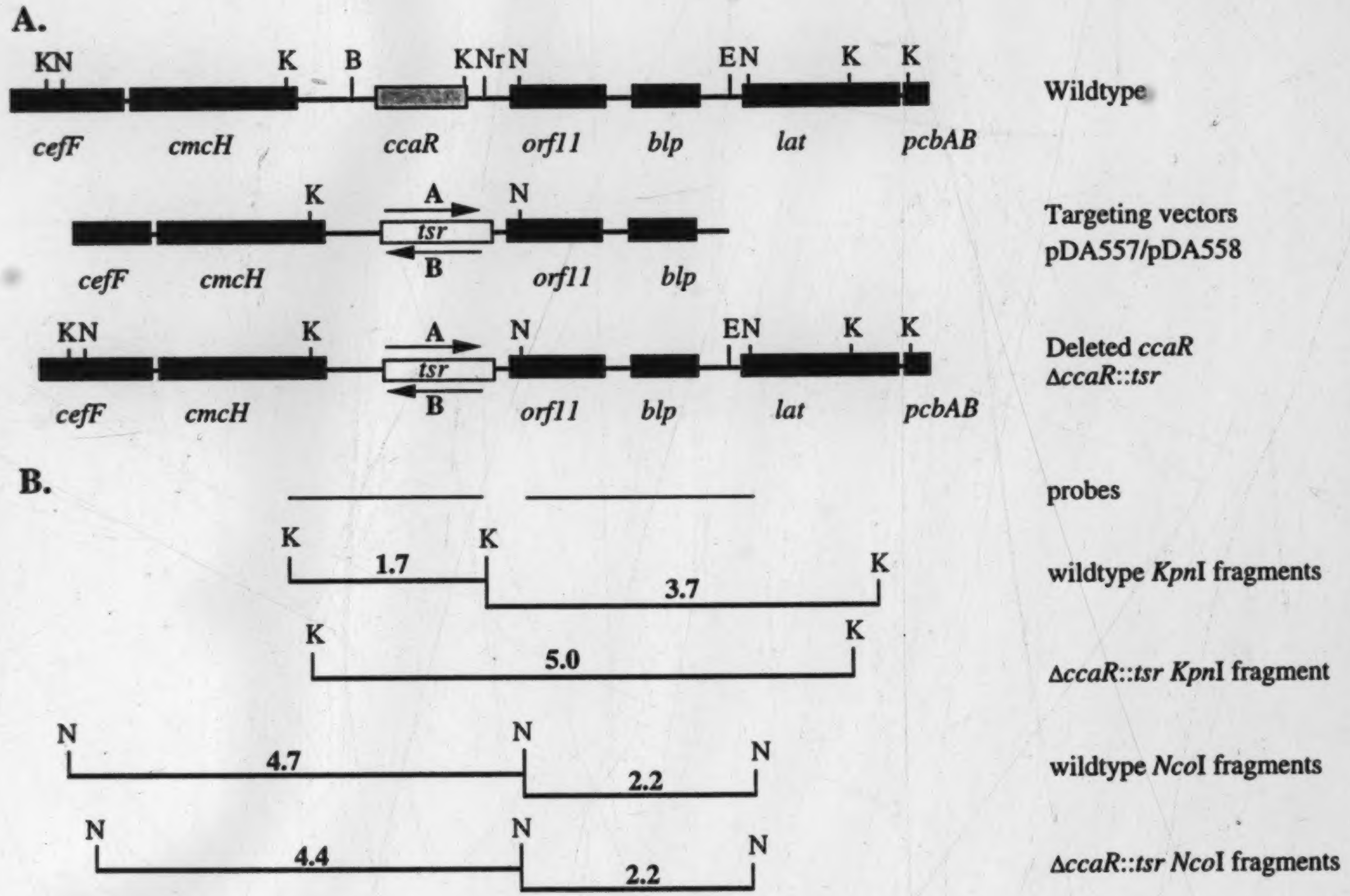

A.
KN
K
B
KNr N
EN
K
K
cefF
cmcH
ccaR
orf11
blp
lat
pcbAB
Wildtype
K
A
tsr
B
N
cefF
cmcH
orf11
blp
Targeting vectors
pDA557/pDA558
KN
K
A
tsr
B
N
EN
K
K
cefF
cmcH
orf11
blp
lat
pcbAB
Deleted ccaR
ΔccaR::tsr
B.
probes
K
1.7
K
3.7
K
wildtype KpnI fragments
K
5.0
K
ΔccaR::tsr KpnI fragment
N
4.7
N
2.2
N
wildtype NcoI fragments
N
4.4
N
2.2
N
ΔccaR::tsr NcoI fragments

Figure 38. Strategy for the deletion of the *cob* genes and replacement with the *tsr* marker. **A.** The dark shaded boxes represent ORFs adjacent to the *cob* genes in the chromosome and the lighter shaded boxes represent the *cob* genes within the wildtype cephamycin cluster. The targeting vectors pDA559 and pDA560 contained a contiguous stretch of *S. clavuligerus* DNA which extends from *cefF* to *pcbAB* except that the 3.7 kb *Bam*HI/*Eco*RI fragment containing the *cob* genes was removed and replaced with the *tsr* marker, represented by the clear box. The *tsr* marker is shown to be present in both orientations (A or B). The mutant chromosome results as the *tsr* marker replaces the wildtype genes in the chromosome. **B.** Southern hybridization pattern of *Kpn*I or *Nco*I digested genomic DNA from wildtype and Δ*cob*::*tsr* mutants using the 1.7 kb *Kpn*I fragment and the 2.2 kb *Nco*I fragment as probes. Abbreviations for restrictions sites: B for *Bam*HI, E for *Eco*RI, K for *Kpn*I and N for *Nco*I.

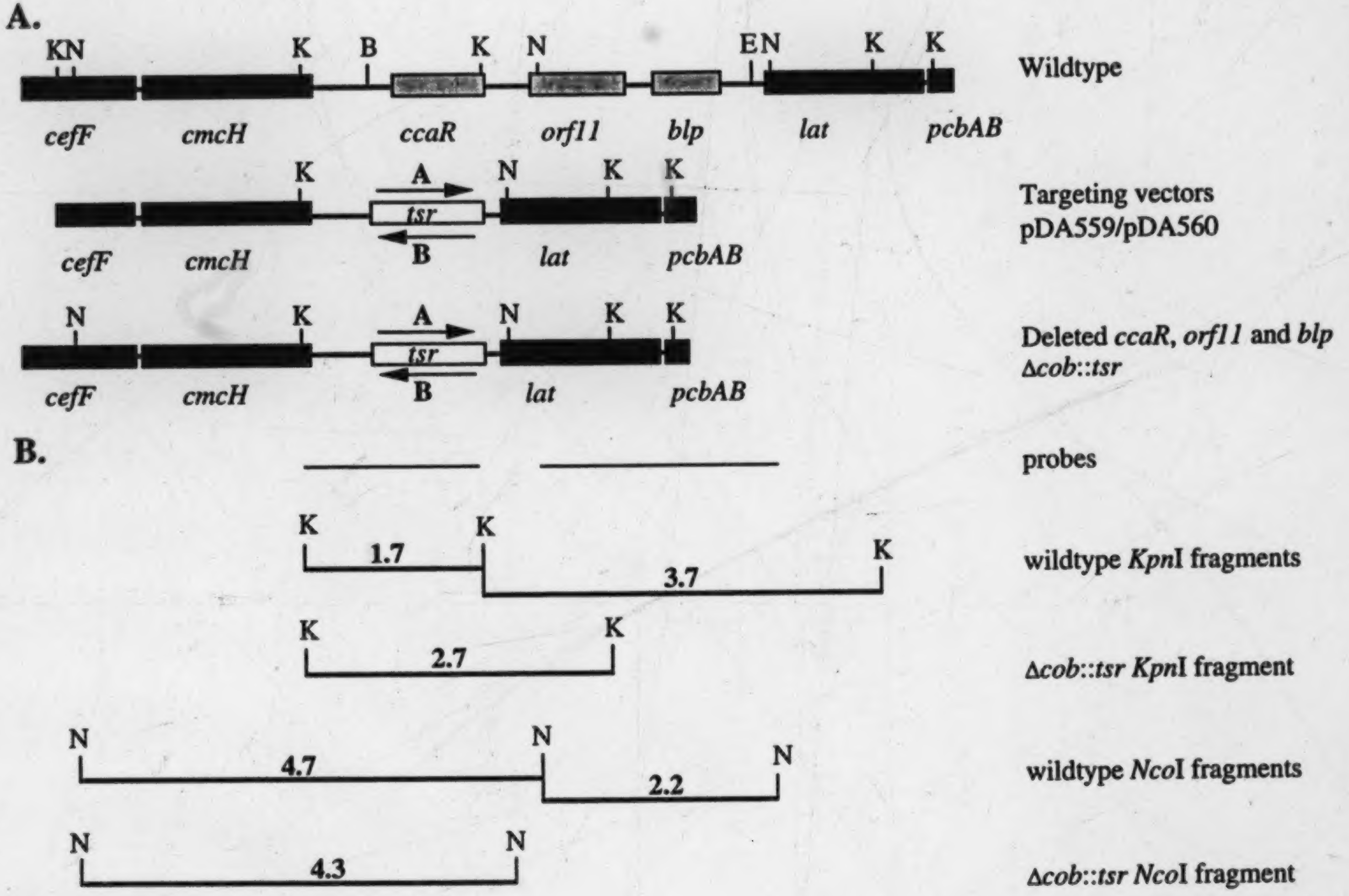
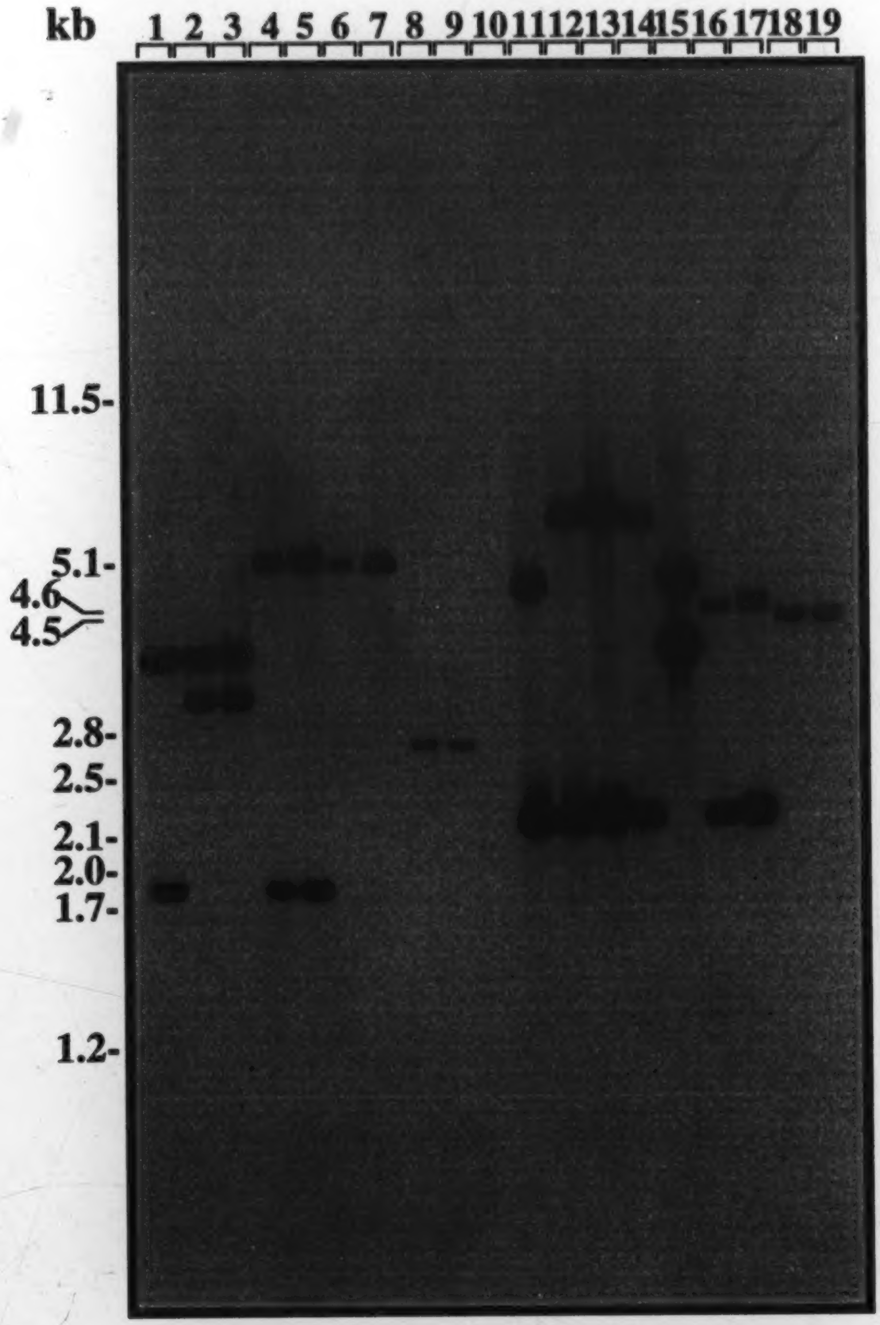

A.
KN
K
B
K
N
EN
K
K
cefF
cmcH
ccaR
orf11
blp
lat
pcbAB
Wildtype
K
A
tsr
B
N
K
K
cefF
cmcH
lat
pcbAB
Targeting vectors
pDA559/pDA560
N
K
A
tsr
B
N
K
K
cefF
cmcH
lat
pcbAB
Deleted ccaR, orf11 and blp
Δcob::tsr
B.
probes
K
1.7
K
3.7
K
wildtype KpnI fragments
K
2.7
K
Δcob::tsr KpnI fragment
N
4.7
N
2.2
N
wildtype NcoI fragments
N
4.3
N
Δcob::tsr NcoI fragment

Figure 39. Confirmation of gene replacement in *ccaR*, *orf11* and *cob* mutants by Southern blot analysis. *Kpn*I or *Nco*I digested genomic DNA isolated from wildtype and *ccaR*, *orf11* and *cob* mutants was hybridized with the 1.7 kb *Kpn*I fragment and the 2.2 kb *Nco*I fragment (see Fig. 35 to 38). *Kpn*I digested genomic DNA (Lanes 1. wildtype *S. clavuligerus*, 2. *ccaR*::*apr*A1 mutant, 3. *ccaR*::*apr*B1 mutant, 4. *orf11*::*apr*A1 mutant, 5. *orf11*::*apr*B1 mutant, 6. Δ*ccaR*::*tsr*A1 mutant, 7. Δ*ccaR*::*tsr*B1 mutant, 8. Δ*cob*::*tsr*A1 mutant and 9. Δ*cob*::*tsr*B1), lane 10. size marker and *Bam*HI digested genomic DNA (Lanes 11. wildtype *S. clavuligerus*, 12. *ccaR*::*apr*A1 mutant, 13. *ccaR*::*apr*B1 mutant, 14. *orf11*::*apr*A1 mutant, 15. *orf11*::*apr*B1 mutant, 16. Δ*ccaR*::*tsr*A1 mutant, 17. Δ*ccaR*::*tsr*B1 mutant, 18. Δ*cob*::*tsr*A1 mutant and 19. Δ*cob*::*tsr*B1) were separated on a 0.7% (w/v) agarose gel.


kb
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
11.5-
5.1-
4.6
4.5
2.8-
2.5-
2.1-
2.0-
1.7-
1.2-

in the absence of antibiotic selection to promote the loss of free plasmid. Putative mutants were selected for maintenance of the tsr^R phenotype but loss of the hyg^R phenotype.

The mutations were confirmed by Southern blot analysis. The digested gDNA was hybridized with the 1.7 kb *Kpn*I fragment and the 2.2 kb *Nco*I fragment (Fig. 38). The wildtype *Kpn*I digested gDNA gave hybridizing bands at 3.7 and 1.7 kb (Fig. 39, lane 1) while the Δ*cob*::*tsr* mutants gave a 2.7 kb hybridizing band (lanes 8-9). The wildtype *Nco*I digested gDNA gave hybridizing bands of 4.7 and 2.2 kb (lane 11), while the Δ*cob*::*tsr* mutants gave a hybridizing band of 4.3 kb (lanes 18-19).

The above results were all consistent with a double crossover event having occurred in the Δ*cob*::*tsr*A and Δ*cob*::*tsr*B mutants of *S. clavuligerus*.

V.1.4 Additional Southern blot analyses of mutant strains

The Southern blots used to confirm the nature of the mutant strains were stripped and reprobed with other fragments of DNA such as the *tsr* marker. The results from these secondary hybridizations are only discussed in the text and reproductions of the autoradiographs are not included.

When the *Kpn*I digested gDNA was hybridized with a DNA fragment containing the *apr* cassette the *ccaR*::*apr* mutants gave a 3.2 kb hybridizing band and the *orf11*::*apr* mutants gave a 5.2 kb hybridizing band. There was no hybridization in the wildtype, Δ*ccaR*::*tsr* mutants or the Δ*cob*::*tsr* mutants lanes. This confirmed the presence of an *apr* cassette within the *ccaR*::*apr* mutants and the *orf11*::*apr* mutants.

When the *Kpn*I digested gDNA was hybridized with the *tsr* marker the Δ*ccaR*::*tsr* mutants gave a 5.0 kb hybridizing band and the Δ*cob*::*tsr* mutants gave a 2.7 kb hybridizing band. There was no hybridization in the wildtype, *ccaR*::*apr* mutants or the *orf11*::*apr* mutants lanes. This confirmed the presence of the *tsr* marker within the Δ*ccaR*::*tsr* mutants and the Δ*cob*::*tsr* mutants.

The *Kpn*I digested gDNA of all the strains did not contain any DNA sequences that hybridized with pJOE829 demonstrating that a double crossover had occurred with removal of the targeting vector.

V.2 Characterization of *ccaR* mutants

The investigation of the mechanisms by which the CcaR transcriptional activator regulates antibiotic biosynthesis required the construction of *ccaR* mutant strains that contained a disrupted or deleted *ccaR* gene but were otherwise isogenic to the wildtype strain. Therefore three different sets of *ccaR* mutants were generated by homologous recombination.

V.2.1 Characterization of the cephamycin C production by the *ccaR* mutants

The four independent mutants (two from each resistance marker orientation) from each of the three sets of *ccaR* mutants (*ccaR*::*apr*, Δ*ccaR*::*tsr* and Δ*cob*::*tsr*) were grown in trypticase soy broth + 1% (w/v) starch (TSBS). Samples were taken at both 48 and 72 hours and analyzed for the production of cephamycin. Cephamycin production was determined by bioassay using the penicillin supersenstive strain *E. coli* ESS (see II.2.4.4.1.).

All three sets of *ccaR* mutants, independent of the antibiotic resistance marker orientation, were unable to make cephamycin even after 72 hours growth in TSBS. In contrast, wildtype *S. clavuligerus* produced significant amounts of cephamycin.

V.2.2 Western blot analysis of the *ccaR* mutants

Cell-extracts were prepared from the *ccaR* mutants and wildtype *S. clavuligerus* strain for western blot analysis. The western blots were developed with antibodies specific for the CcaR protein as well as the LAT, IPNS and DAOCS biosynthetic enzymes.

Western blot analysis was preferred as the method to characterize the mutants rather than enzyme assays. As well, the CcaR protein presently does not have a biochemical assay to test for its function and therefore could only be analyzed by western blot analysis. The quality of some western blots developed with CcaR-specific antibodies were poor because of non-specific cross-reactivity with other proteins and possibly the low amount of the CcaR protein in cell-extracts.

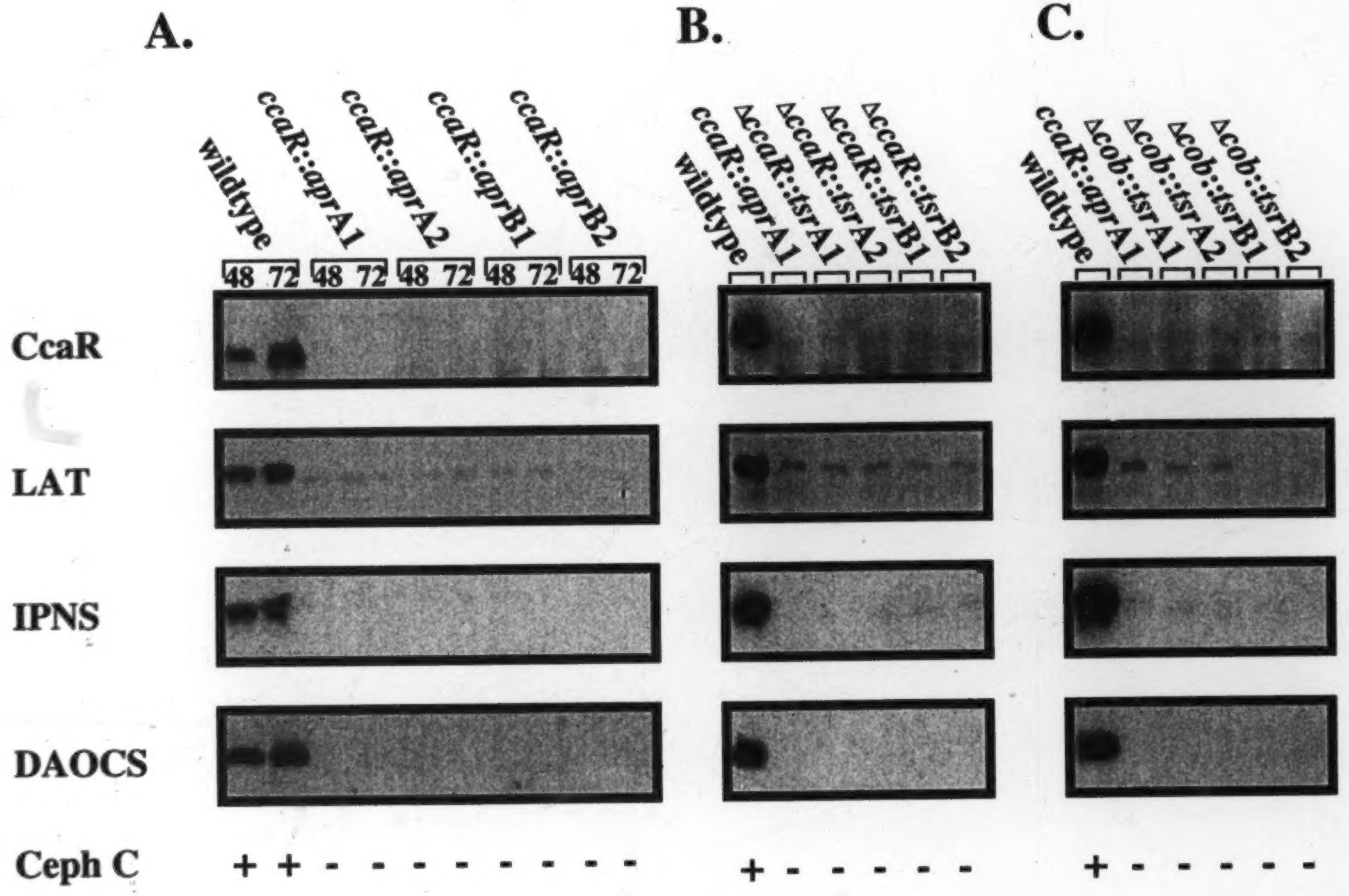
V.2.2.1 Western blot analysis of the *ccaR*::*apr* mutants

The *ccaR*::*apr* mutants as expected did not make any CcaR protein after 48 or 72 hours growth when compared to wildtype *S. clavuligerus* (Fig. 40A). They also did not produce any or had greatly decreased amounts of the LAT protein when compared to wildtype *S. clavuligerus*. These four mutants also did not produce detectable amounts of the IPNS protein or the DAOCS protein, while high levels of protein were present in the wildtype *S. clavuligerus* lanes.

V.2.2.2 Western blot analysis of the Δ*ccaR*::*tsr* mutants

The Δ*ccaR*::*tsr* mutants as expected did not make any CcaR protein when compared to the wildtype *S. clavuligerus* (Fig. 40B). They also did not produce any of

Figure 40. Presence of the CcaR protein and the cephamycin biosynthetic enzymes in *ccaR* mutant strains as detected by western analysis. **A.** Wildtype and *ccaR*::*apr* disruption strains. **B.** Wildtype and Δ*ccaR*::*tsr* deletions strains. **C.** Wildtype and Δ*cob*::*tsr* deletion strains. Approximately 10 μg of cell-extract protein from each strain, harvested after 72 hours growth unless otherwise noted, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the CcaR, LAT, IPNS and DAOCS proteins. The ability of each strain to produce cephamycin C (Ceph C) was determined by bioassay and the results were scored as: + for production and - for lack of production.


A.
B.
C.
wildtype
ccaR::aprA1
ccaR::aprA2
ccaR::aprB1
ccaR::aprB2
48 72
48 72
48 72
48 72
48 72
wildtype
ccaR::aprA1
∆ccaR::tsrA1
∆ccaR::tsrA2
∆ccaR::tsrB1
∆ccaR::tsrB2
wildtype
ccaR::aprA1
∆cob::tsrA1
∆cob::tsrA2
∆cob::tsrB1
∆cob::tsrB2
CcaR
LAT
IPNS
DAOCS
Ceph C
+ + - - - - - - - -
+ - - - - -
+ - - - - -

the LAT, IPNS or DAOCS biosynthetic enzymes necessary for the biosynthesis of cephamycin. In contrast, wildtype *S. clavuligerus* which did make cephamycin C had high levels of these proteins.

V.2.2.3 Western blot analysis of the Δ*cob*::*tsr* mutants

The Δ*cob*::*tsr* mutants as expected did not make any CcaR protein when compared to the wildtype *S. clavuligerus* (Fig. 40C). They also did not produce any of the LAT, IPNS or DAOCS biosynthetic enzymes but wildtype *S. clavuligerus* had high levels of these proteins.

Inactivation of the *ccaR* gene or deletion of the *ccaR* gene or the *cob* genes eliminates production of the CcaR protein and dramatically decreases or eliminates production of the LAT, IPNS and DAOCS biosynthetic proteins. It was not determined whether the small amount of protein detected by the LAT-specific antibodies was due to cross-reactivity or the continued production of low levels of LAT protein.

V.3 Characterization of the NTG mutants

A number of cephamycin non-producing mutants derived from N-methyl-N'-nitro-N-nitrosoguanidine (NTG) mutagenesis on wildtype *S. clavuligerus* spores were available from previous studies (169). Preliminary characterization of the mutants suggested that some were unable to make the polycistronic message originating from the *lat* promoter as well as some of the biosynthetic enzyme activities (177). Four NTG mutants (19, 26, 32 and 83) were further characterized and confirmed to be cephamycin non-producers and when pairs of these strains were co-fermented they were still unable to produce cephamycin (data not shown). The inability of these NTG mutants to co-synthesize suggested that the mutations could be regulatory and the mutants were unable to co-synthesize because they lacked a number of biosynthetic enzyme activities, although other possibilities could explain the inability of these mutants to co-synthesize. The mutants may all possess mutations of the same gene and therefore lack an identical enzymatic step or more likely they may have multiple mutations or large deletions removing the genes for a number of biosynthetic enzymes.

The western analysis system used to characterize the *ccaR* mutants created in this study allowed the NTG mutants to be further characterized. With western analysis it would be possible to determine if the NTG mutants simply lacked a certain biosynthetic enzyme or if they behaved like a *ccaR* mutant and lacked the CcaR protein as well as all the biosynthetic enzymes tested.

V.3.1 Production of cephamycin C by the NTG mutants

The four NTG mutants were grown in TSBS along with wildtype *S. clavuligerus* and Δ*ccaR*::*tsr*A1 mutant controls. Samples were taken at both 48 and 72 hours and analyzed for the production of cephamycin.

All of the NTG mutants as well as the Δ*ccaR*::*tsr*A1 mutant were unable to make cephamycin even after 72 hours growth in TSBS. In contrast, wildtype *S. clavuligerus* produced significant amounts of cephamycin.

V.3.2 Western blot analysis of the NTG mutants

Cell-extracts were prepared from the NTG mutants, the Δ*ccaR*::*tsr*A1 mutant and the wildtype control strains for western blot analysis. The western blots were developed with antibodies specific for the CcaR protein as well as the LAT, IPNS and DAOCS biosynthetic enzymes.

When examined by western analysis the Δ*ccaR*::*tsr*A1 mutant as expected did not make any CcaR protein or any of the biosynthetic enzymes when compared to the wildtype strain (Fig. 41). Three of the four NTG mutants; NTG 26, NTG 32 and NTG 83 behaved like the Δ*ccaR*::*tsr*A1 mutant and lacked production of the CcaR protein. These mutants also lacked production of the LAT, IPNS and DAOCS biosynthetic enzymes. The NTG 19 mutant behaved like wildtype *S. clavuligerus* producing the CcaR protein and the biosynthetic enzymes but was still unable to make cephamycin. Previously, the NTG 19 mutant was characterized and shown to contain an uncharacterized mutation in the ACVS protein which still allowed the extremely large protein to be produced but the enzyme was determined to be inactive (166).

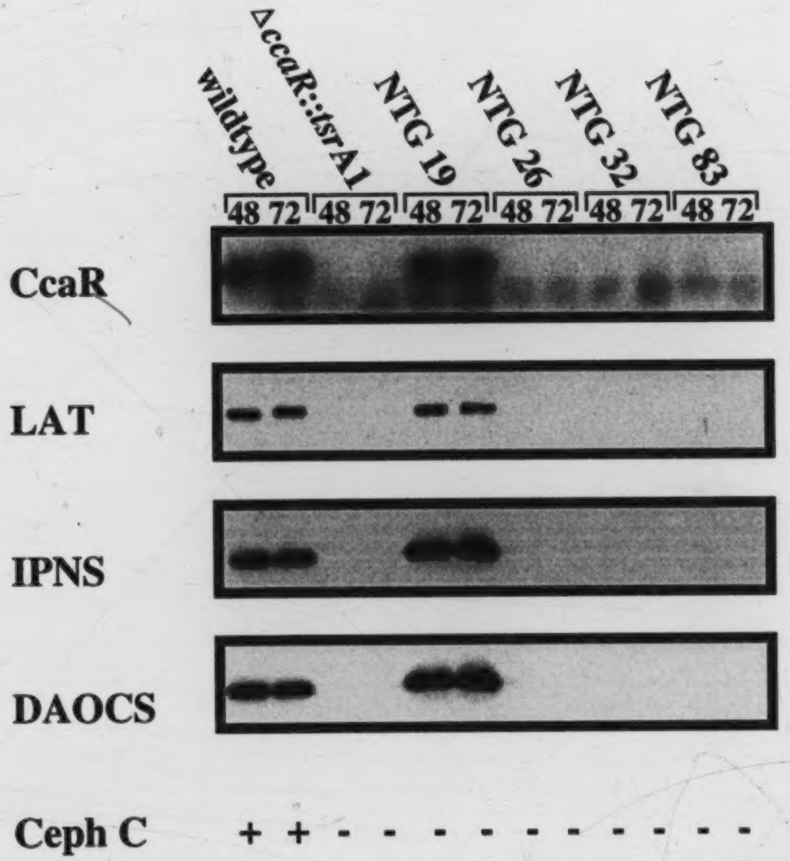
V.4 Comparison of cephamycin biosynthetic activities in the various *ccaR* mutants

A single representative from each set of the *ccaR* mutants, the NTG mutants showing the *ccaR* mutant phenotype and the wildtype strain were all grown in TSBS for 48 hours. The cultures were split in half and the harvested cell pellets were sonicated in ACVS or TDE breakage buffer and assayed for certain cephamycin biosynthetic enzyme activities as well as being examined by western blot analysis and SDS-PAGE.

V.4.1 Biochemical characterization of the *ccaR* mutants

The cell-extracts prepared in ACVS breakage buffer were assayed for ACVS enzyme activity (see II.2.4.2.1). The cell-extracts prepared in TDE breakage buffer were

Figure 41. Presence of the CcaR protein and cephamycin biosynthetic enzymes in the NTG mutant strains as detected by western analysis. Approximately 10 µg of cell-extract protein from each strain, harvested after 72 hours growth, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the CcaR, LAT, IPNS and DAOCS proteins. The ability of each strain to produce cephamycin C (Ceph C) was determined by bioassay and the results were scored as: + for production and - for lack of production.


wildtype
ΔccaR::tsrA1
NTG 19
NTG 26
NTG 32
NTG 83
48 72 48 72 48 72 48 72 48 72 48 72
CcaR
LAT
IPNS
DAOCS
Ceph C
+ + - - - - - - - - - -

used in enzyme assays measuring IPNS, IPNE or DAOCS activity (see II.2.4.2.2, II.2.4.2.3 and II.2.4.2.4).

The *ccaR* mutants did not make any ACVS enzyme activity nor did the NTG mutants, whereas wildtype *S. clavuligerus* had significant amounts of ACVS activity. The *ccaR* mutants were also unable to make any of the IPNS, IPNE and DAOCS biosynthetic enzyme activities all of which were easily detected in wildtype *S. clavuligerus*. The IPNS, IPNE and DAOCS biosynthetic enzyme activities were also absent from the NTG mutants.

V.4.2 Western blot analysis of the *ccaR* mutants

The cell-extracts prepared in TDE buffer from the NTG mutants, the *ccaR* mutants and the wildtype control were subjected to western blot analysis. The western blots were developed with antibodies specific for the CcaR protein as well as the LAT, IPNS and DAOCS biosynthetic enzymes.

When examined by western analysis the *ccaR* mutants and the NTG mutants did not make any CcaR protein (Fig. 42A). All of the *ccaR* mutants and the NTG mutants do not make the LAT, IPNS and DAOCS biosynthetic enzymes. The IPNS and DAOCS enzyme activities were already shown to be absent from these mutant strains. The results generated from the western analysis were identical to the enzyme assay results.

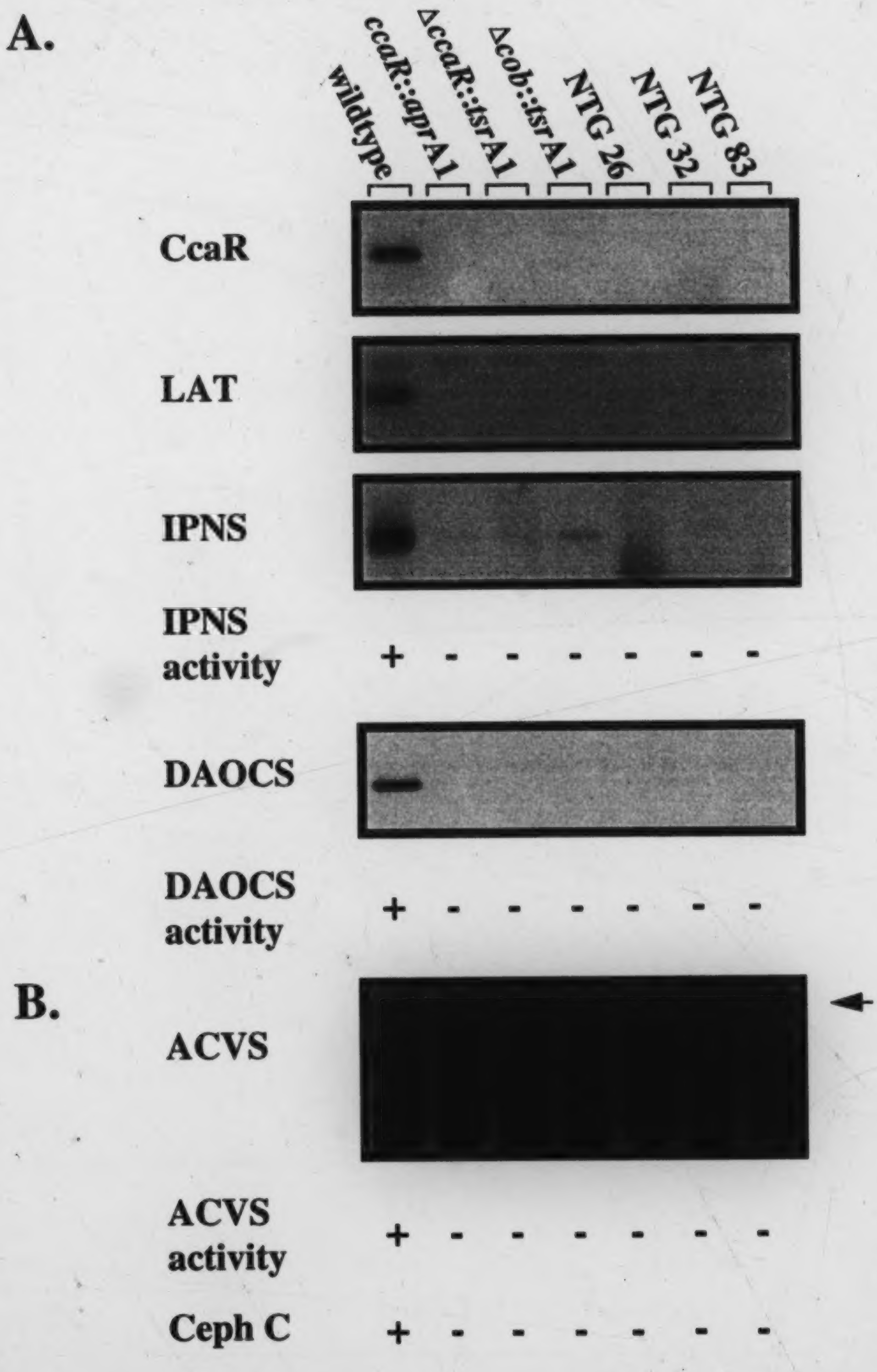
V.4.3 Analysis of the *ccaR* mutants for production of ACVS by SDS-PAGE

The cell-extracts prepared in TDE buffer from the NTG mutants, the *ccaR* mutants and the wildtype control were analyzed by SDS-PAGE. Again the results observed from the SDS-PAGE were identical to the enzyme assays for the mutants. The ACVS enzyme assay suggested that the mutants did not have any ACVS activity and the SDS-PAGE confirmed that the mutants did not produce detectable levels of the ACVS protein (Fig. 42B).

V.5. Characterization of the *orf11::apr* mutants

The *orf11* gene and its role in antibiotic production was also investigated. The *orf11::apr* mutants resulted from the insertion of the *apr* cassette into the unique *Nco*I site within the *orf11* coding region (see V.1.2.2).

Figure 42. Presence of the CcaR protein and the cephamycin biosynthetic enzymes in the *ccaR* and NTG mutant strains. **A.** Approximately 10 μg of cell-extract protein from each strain, harvested after 48 hours growth, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the CcaR, LAT, IPNS and DAOCS proteins. **B.** Approximately 75 μg of cell-extract protein from each strain was separated by 10% SDS-PAGE and stained with coomassie blue. The ability of each strain to produce cephamycin C (Ceph C) was determined by bioassay and the ability of each strain to produce the certain cephamycin biosynthetic activities was determined by enzyme assay the results were scored as: + for production and - for lack of production.


A.
wildtype
ccaR::aprA1
ΔccaR::tsrA1
Δcob::tsrA1
NTG 26
NTG 32
NTG 83
CcaR
LAT
IPNS
IPNS activity + - - - - - -
DAOCS
DAOCS activity + - - - - - -
B.
ACVS
ACVS activity + - - - - - -
Ceph C + - - - - - -

V.5.1 Production of cephamycin C by *orf11::apr* mutants

The four independent mutants (two from each resistance marker orientation) from the *orf11::apr* mutants were grown in TSBS. Samples were taken at both 48 and 72 hours and analyzed for cephamycin production.

Both the *orf11::apr*A and *orf11::apr*B mutants were capable of producing cephamycin. The *orf11::apr*A mutants produced an average of 92% and 138% of wildtype levels after 48 and 72 hours, respectively. The *orf11::apr*B mutants produced an average of 16% and 33% of wildtype levels after 48 and 72 hours, respectively. Both *orf11::apr*B mutants displayed this decrease in cephamycin production in two separate fermentations.

Mutation of the *orf11* gene did not eliminate cephamycin production but the *orf11::apr*B mutants produced much lower levels of cephamycin than wildtype *S. clavuligerus*. However, the *orf11::apr*A mutants continue to make near wildtype levels of cephamycin C suggesting that the disruption of *orf11* gene is not the cause for the decreased cephamycin production.

V.6 Production of other β-lactam antibiotics by mutants strains

The TSBS medium was routinely used in the characterization of mutant strains for the ability to produce cephamycin, however this medium only supports high level production of cephamycin but not clavulanic acid or clavams. The high level production of clavulanic acid and clavams, the other β-lactam metabolites produced by *S. clavuligerus*, requires growth in specific media and this media-specific production has been tracked to the level of transcription (168). *S. clavuligerus* grown in the defined starch asparagine (SA) minimal medium is capable of high level production of both clavulanic acid and cephamycin (4). This medium does not support the production of clavams and the *cas1* gene within the clavam biosynthetic cluster has been shown to be transcriptionally silent in SA media (168). Growth of *S. clavuligerus* in the complex SOY medium (192) results in high levels of clavulanic acid, cephamycin and clavams being produced.

V.6.1 Characterization of β-lactam production by mutant strains

A single representative from each set of *ccaR* mutants, the NTG mutants, the *orf11::apr*A1 mutant and the wildtype *S. clavuligerus* control were all grown in each of these media. Samples were taken at 72 hours and analyzed for the production of cephamycin (all three media), clavulanic acid (SA and SOY media) and clavams (SOY medium only). Production of clavulanic acid and clavams was determined by HPLC analysis and compared to authentic standards (see II.2.4.4.2). The results of the β-lactam antibiotic production were summarized in Table 16.

All of the *ccaR* mutants and NTG mutants were unable to produce cephamycin on any of the media tested. These same media supported high level production of cephamycin by wildtype *S. clavuligerus*. The *ccaR* mutants and the NTG 26, NTG 32 and NTG 83 mutants were also unable to produce clavulanic acid or clavams when grown in SA or SOY media, where as wildtype *S. clavuligerus* was capable of producing all of these metabolites. The NTG 19 strain was capable of producing clavulanic acid and clavams but not cephamycin, confirming that NTG 19 was a cephamycin non-producer because of the ACVS mutation and not a regulatory mutant. The *orf11*::*apr*A1 strain was capable of producing high levels of cephamycin, clavulanic acid and clavams.

HPLC analysis of supernatants from wildtype *S. clavuligerus* grown in SOY medium shows peaks for: clavaminic acid, a precursor of both clavulanic acid and clavams, clavulanic acid, as well as two of the clavams, hydroxymethyl clavam and clavam-2-carboxylate (Fig. 43A). The *ccaR* mutants (*ccaR*::*apr*A1 and Δ*ccaR*::*tsr*A1) were unable to make any of these metabolites detected by HPLC analysis (Fig. 43 B and C). A similar HPLC profile was observed for the *orf2*::*apr* mutants grown in SOY media (67). The *orf2* gene is located at one end of the clavulanic acid gene cluster and is believed to be one of the first steps in biosynthesis and therefore is essential for both clavulanic acid and clavam biosynthesis.

V.7 Genetic characterization of the *ccaR* mutants generated by random mutagenesis

During the characterization of their clavulanic acid non-producing mutants Barton and coworkers (20, 221) found two classes of mutants which were unable to crossfeed with other *dcl* mutants. Unexpectedly, these mutants were also unable to make cephamycin. Characterization of the *dclZ*124 strain revealed a large deletion removing the *dclX* (*ccaR*) gene as well as some of the *orf11* and *cmcH* genes. Characterization of the *dclX*68 strain did not reveal any changes in the chromosomal organization and therefore it was felt that the *dclX*68 strain contained a point mutation within the *ccaR* gene.

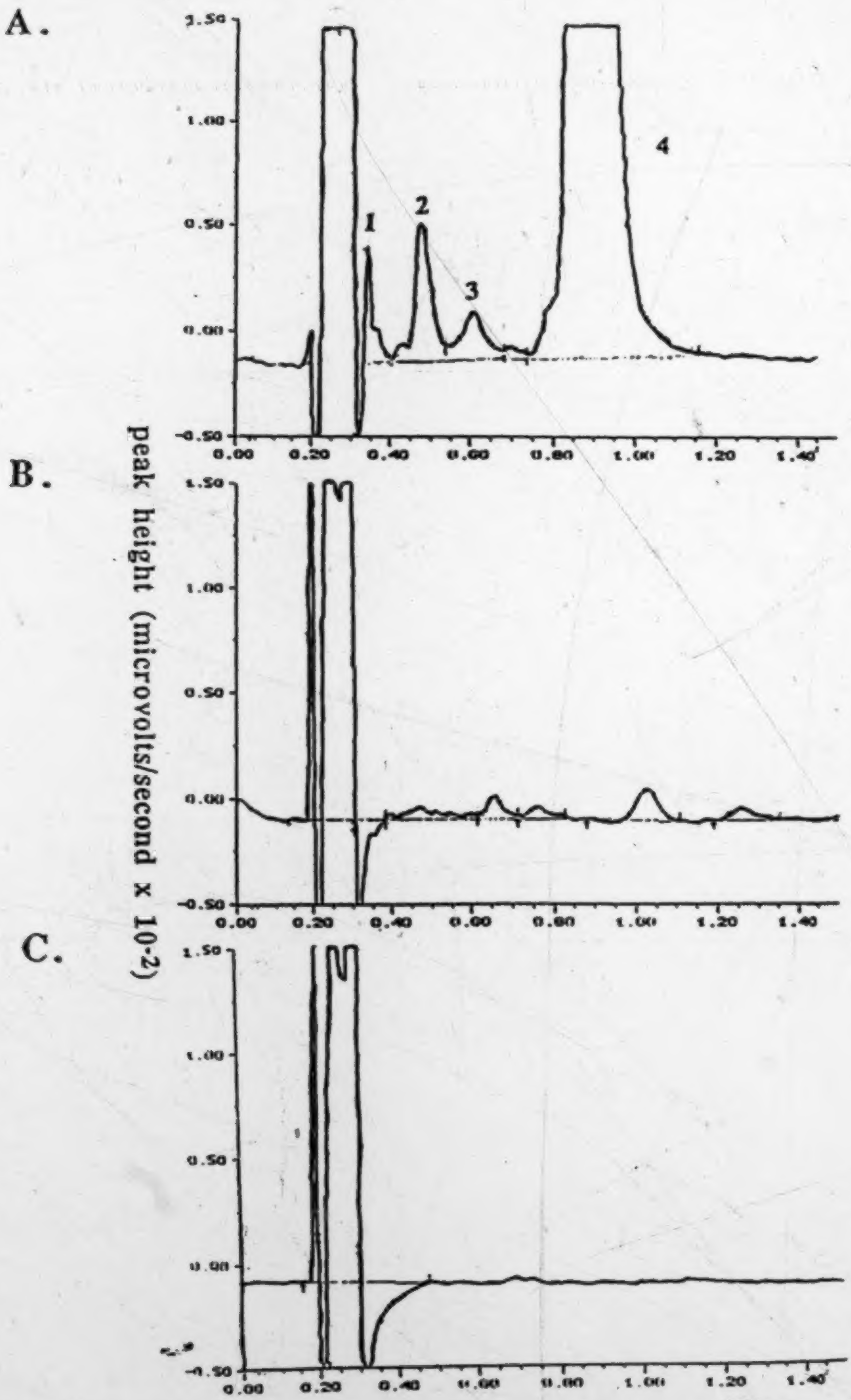
The nature of the mutations in the *dclX*68, NTG 26, NTG 32 and NTG 83 mutants was further characterized because these particular mutants did not produce the CcaR protein, lacked production of certain cephamycin biosynthetic enzymes and did not produce cephamycin or clavulanic acid, and therefore fit the regulatory mutant phenotype. The *dclX*68 mutant was not directly tested for its ability to produce cephamycin enzymes but the mutant was demonstrated to be unable to produce α-aminoadipate (20). On the

Table 16. Production of β-lactam antibiotics by *ccaR*, *cob* and NTG mutants

	TSBS	Starch Asparagine		SOY		
Strain	Cephamycin C	Cephamycin C	Clavulanic acid	Cephamycin C	Clavulanic acid	Clavams[a]
wildtype	+	+	+	+	+	+
ccaR::*apr*A1	-	-	-	-	-	-
ccaR::*tsr*A1	-	-	-	-	-	-
cob::*tsr*A1	-	-	-	-	-	-
orf11::*apr*A1	+	+	+	+	+	+
NTG 19	-	-	+	-	+	+
NTG 26	-	-	-	-	-	-
NTG 32	-	-	-	-	-	-
NTG 83	-	-	-	-	-	-

[a] Clavams includes both carboxy-2-clavam and hydroxymethyl-clavam and both metabolites are detected by HPLC analysis.

Figure 43. HPLC analysis of clavulanic acid and clavam metabolite production by cultures grown in SOY fermentation medium. **A.** Wildtype *S. clavuligerus* culture supernatant (diluted 1:5). **B.** *ccaR*::*apr* mutant culture supernatant (diluted 1:2). **C.** Δ*ccaR*::*tsr* mutant culture supernatant (diluted 1:2). **1.** clavaminic acid, **2.** clavam-2-carboxylate, **3.** 2-hydroxymethylclavam, **4.** clavulanic acid.


A.
B.
C.
1
2
3
4
1.50
1.00
0.50
0.00
-0.50
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
peak height (microvolts/second x 10-2)
retention times (minutes x 10-1)

other hand, the NTG 19 mutant produced the CcaR protein, the cephamycin biosynthetic enzymes and clavulanic acid and therefore was not further characterized.

V.7.1 Southern blot analysis of NTG mutants

Although the actual *dcl* mutant strains were not available it was possible to obtain gDNA from the *dclX*68 mutant. Samples of gDNA from the mutants *dclX*68, NTG 26, NTG 32 and NTG 83 were digested, along with wildtype gDNA with *Kpn*I or *Sal*I for Southern blot analysis. After transfer of the digested DNA to a nylon membrane, the membrane was hybridized with a 1.7 kb *Kpn*I fragment containing almost the entire *ccaR* gene and some of the *cmcH* gene (data not shown). Hybridization with this fragment to the wildtype gDNA would give a 1.7 kb band in the *Kpn*I digests and two bands of 1.1 kb and 1.8 kb for the *Sal*I digests. Bands of these sizes were observed in the wildtype lanes and similar sized bands were seen in the *Kpn*I or *Sal*I digested gDNA from the NTG mutants and *dclX*68 mutant strain.

V.7.2 Sequence analysis of the *ccaR* mutants alleles

Since none of the mutant strains appear to have a sizable deletion in the *ccaR* gene region they were presumed to have point mutations. Two methods could be used to determine the sequence of the *ccaR* alleles from these mutant strains. PCR could be used to amplify the *ccaR* gene from the gDNA of each mutant and the amplified product could be sequenced, or a DNA fragment containing the *ccaR* gene could be cloned from the gDNA of each strain and the sequence of this fragment could be determined. So that the different *ccaR* alleles could be used in complementation experiments it was decided that the *Bam*HI/*Nru*I fragments containing the *ccaR* allele from each mutant would be inserted into the pKS(+) vector.

The gDNA purified from each strain was digested with *Bam*HI and then *Nru*I, separated on an agarose gel and agarose gel blocks containing DNA of approximately 1.0 to 1.8 kb in size (the *ccaR*-containing *Bam*HI/*Nru*I fragment would be 1.4 kb) were removed from the gel. Successful gel purification of the *ccaR* containing *Bam*HI/*Nru*I fragment was confirmed by Southern blot analysis (data not shown). The purified *Bam*HI/*Nru*I fragments were ligated with pKS(+) digested with *Bam*HI and *Eco*RV and then dephosphorylated. The resulting ligation mixtures were first digested with *Pst*I to remove any religated vector and then *Nar*I to remove some of the non-*ccaR*-containing *Bam*HI/*Nru*I fragments cloned. The *Nar*I recognition site (GG/CGCC) occurs frequently in the high GC *Streptomyces* DNA. In fact, there are 42 *Nar*I sites within the 26 kb of sequence compiled in this study but none of these *Nar*I sites are present within the 1.4 kb

*Bam*HI/*Nru*I fragment containing the *ccaR* gene. The transformed ligation mixtures were plated onto 2x YT plates supplemented with ampicillin and glucose and incubated until colonies appeared. The colonies were replica plated onto duplicate plates, one of which contained a nylon filter and positive clones were detected by colony hybridization using a *ccaR*-specific probe. A single positive clone from each mutant was subjected to sequence analysis and compared to the wildtype *ccaR* sequence (Fig. 44).

The NTG 83 mutant had a single base change (C to T) that results in a proline converting to a serine at amino acid 55 but otherwise creates a full length unaltered protein. The *dclX*68 mutant contained a single base change (G to A) that generates a stop codon creating a truncated protein of 66 amino acids. The NTG 32 mutant had a deletion of a single base causing a frameshift and the generation of a mutant CcaR protein composed of 182 amino acids of authentic CcaR protein and 14 amino acids of missence sequence before being prematurely terminated. Surprisingly, the NTG 26 mutant allele did not have any mutations. None of the NTG mutant *ccaR* alleles contained any mutations within the promoter sequence upstream of the *ccaR* gene. These mutations suggest that the C-terminal end of the CcaR protein is important for function of the protein.

V.8 Complementation of *ccaR* mutants

The *dclX* (*ccaR*) gene was originally discovered because of its ability to restore production of clavulanic acid and cephamycin when supplied *in trans* to the *dclX*68 mutant (221). Complementation of the *ccaR* mutants created and characterized in this study was of interest for a number of reasons. Firstly, addition of the *ccaR* gene *in trans* would confirm that the lack of antibiotic production by these mutants was due to the mutation of the *ccaR* gene. As well, modifications to the *ccaR* gene and their effect on the regulation of antibiotic biosynthesis could be studied.

Complementation studies relied on the pSET152-based plasmids which integrate into the genome of *S. clavuligerus* (27, 126). Verification of plasmid integration was confirmed by genomic PCR (see IV.7.1). The amplified PCR product from gDNA of the complemented strain was compared to the product amplified from free plasmid isolated from *E. coli*.

V.8.1 Complementation of Δ*cob*::*tsr* mutant

Since antibiotic production was lost in all *ccaR*::*apr* and Δ*ccaR*::*tsr* mutants regardless of the orientation of the disruption marker it was concluded that a functional *ccaR* gene was essential for antibiotic production. Therefore the loss of antibiotic

Figure 44. Sequence determination of the *ccaR* alleles from *ccaR* mutants. The 1.4 kb *Bam*HI/*Nru*I fragments cloned from *dclX*68, NTG 26, NTG 32 and NTG 83 genomic DNA were sequenced and the sequences of each *ccaR* allele were compared to the wildtype *ccaR* sequence of Walters et al. (221). The sequence of the mutant alleles was only presented in the region that contained the mutation. The site of the change in DNA sequence is underlined and in the case of the single base deletion for NTG 32 the bases adjacent to the deletion are also underlined. The mutant amino acid(s) are present in lower case and stop codons are represented by asterisks.

```
          1                             11
Wt        GTGACGATCCGGCTCCTGGGGCCGGTGACACTCGTGAAAGGTTCCGTACCGATACCCATC
Seq       V  T  I  R  L  L  G  P  V  T  L  V  K  G  S  V  P  I  P  I

          21                            31
Wt        CGCGGGCAGCGACAGCGGCGATTCCTCGCCTCATTAGCGCTGCGACCGGGCCAGGTCATC
Seq       R  G  Q  R  Q  R  R  F  L  A  S  L  A  L  R  P  G  Q  V  I

          41                            51
Wt        TCCAAGGAAGCGATCATCGAAGACTCCTGGGACGGGGAGCCACCACTGACCGTTTCGGGC
Seq       S  K  E  A  I  I  E  D  S  W  D  G  E  P  P  L  T  V  S  G
NTG 83    TCCAAGGAAGCGATCATCGAAGACTCCTGGGACGGGGAGCCATCACTGACCGTTTCGGGC
Seq       S  K  E  A  I  I  E  D  S  W  D  G  E  P  s  L  T  V  S  G

          61                            71
Wt        CAGTTGCAGACGTCGGCCTGGATGATCCGGACCGCGCTGGCGGAGGCGGGGCTGCCCCGC
Seq       Q  L  Q  T  S  A  W  M  I  R  T  A  L  A  E  A  G  L  P  R
dclX68    CAGTTGCAGACGTCGGCCTAG
Seq       Q  L  Q  T  S  A  *

          81                            91
Wt        GACGCCCTCGGCTCCCACGACCGCGGCTACGAACTGCGCGTCCTGCCGGACTCCATCGAC
Seq       D  A  L  G  S  H  D  R  G  Y  E  L  R  V  L  P  D  S  I  D

          101                           111
Wt        CTCTTCGTCTTCCGGGAGGCCGTGCGCGCCGTGCGGGACCTGCACGCACGCGGTCAGCAC
Seq       L  F  V  F  R  E  A  V  R  A  V  R  D  L  H  A  R  G  Q  H

          121                           131
Wt        CAGGAGGCGTCCGAACGGCTCGACACGGCGCTCGCCCTGTGGAAGGGGCCCGCCTTCGCG
Seq       Q  E  A  S  E  R  L  D  T  A  L  A  L  W  K  G  P  A  F  A

          141                           151
Wt        GATGTGACCTCCAGTCGGCTGCGGCTGCGGGGCGAGACCCTGGAGGAGGAGCGGACCGCC
Seq       D  V  T  S  S  R  L  R  L  R  G  E  T  L  E  E  E  R  T  A

          161                           171
Wt        GCGGTCGAGCTGCGCGCCCTGATCGATGTCGGCCTCGGCTACTACGGGGACGCGATCACC
Seq       A  V  E  L  R  A  L  I  D  V  G  L  G  Y  Y  G  D  A  I  T

          181                           191
Wt        CGGCTGTCGGAGCTCGTCGATCACGACCCGTTCCGTGAGGACCTGTATGTGAGCCTGATG
Seq       R  L  S  E  L  V  D  H  D  P  F  R  E  D  L  Y  V  S  L  M
NTG 32    CGGCTGCGGAGCTCGTCGATCACGACCCGTTCCGTGAGGACCTGTATGTGA
Seq       R  L  r  s  s  s  i  t  t  r  s  v  r  t  c  m  *

          201                           211
Wt        AAGGCCTACTACGCGGAGGGCCGCCAGGCCGACGCGATCCAGGTCTTCCACCGCGCGAAG
Seq       K  A  Y  Y  A  E  G  R  Q  A  D  A  I  Q  V  F  H  R  A  K

          221                           231
Wt        GACATCCTGCGGGAGCAGATCGGCATCAGCCCCGGCGAGCGGATGACAAGGGTCATGCAG
Seq       D  I  L  R  E  Q  I  G  I  S  P  G  E  R  M  T  R  V  M  Q

          241                           251
Wt        GCCATCCTGCGTCAGGACGAGCAGGTCCTGCGGGTCGGTACCCCGGCCTGA
Seq       A  I  L  R  Q  D  E  Q  V  L  R  V  G  T  P  A  *
```

production in the Δ*cob*::*tsr* mutants was predictable due to the lack of a functional *ccaR* gene. Any effects of the mutations in *orf11* or *blp* genes on antibiotic production were obscured in the Δ*cob*::*tsr* mutants by the overriding effect of the *ccaR* mutation. Complementation of the Δ*cob*::*tsr* mutant would make it possible to determine the role of each gene in the regulation of antibiotic biosynthesis without being concerned about polar mutations affecting the results. The insertion and excision of the *apr* cassette made it possible to create three different *cob* DNA fragments, each with one of the *cob* genes specifically mutated by insertion of the translational stop fragment (TSF). Since the TSF is only 37 bp it seems unlikely that it would significantly alter transcription.

It was not possible to create a *blp*::*apr* mutant by homologous recombination in a parental background capable of producing antibiotic so the role of the Blp protein in regulation of antibiotic production could not be determined directly. With the Δ*cob*::*tsr* mutant it was possible to determine if Blp played a role in antibiotic production by complementing the Δ*cob*::*tsr* mutant with a *cob* DNA fragment that had the *blp* gene mutated.

The Δ*cob*::*tsr*A1 mutant was transformed with pSET152-based vectors (Fig. 45) containing the wildtype *cob* genes (pDA1001), the *ccaR* mutant fragment (TSF in *Eco*ICRI near 3' end; pDA1002), the *orf11* mutant fragment (TSF in *Nco*I; pDA1003) and *blp* mutant fragment (TSF in *Not*I; pDA1004) or the negative control plasmid containing the *tsr* marker (pDA1000).

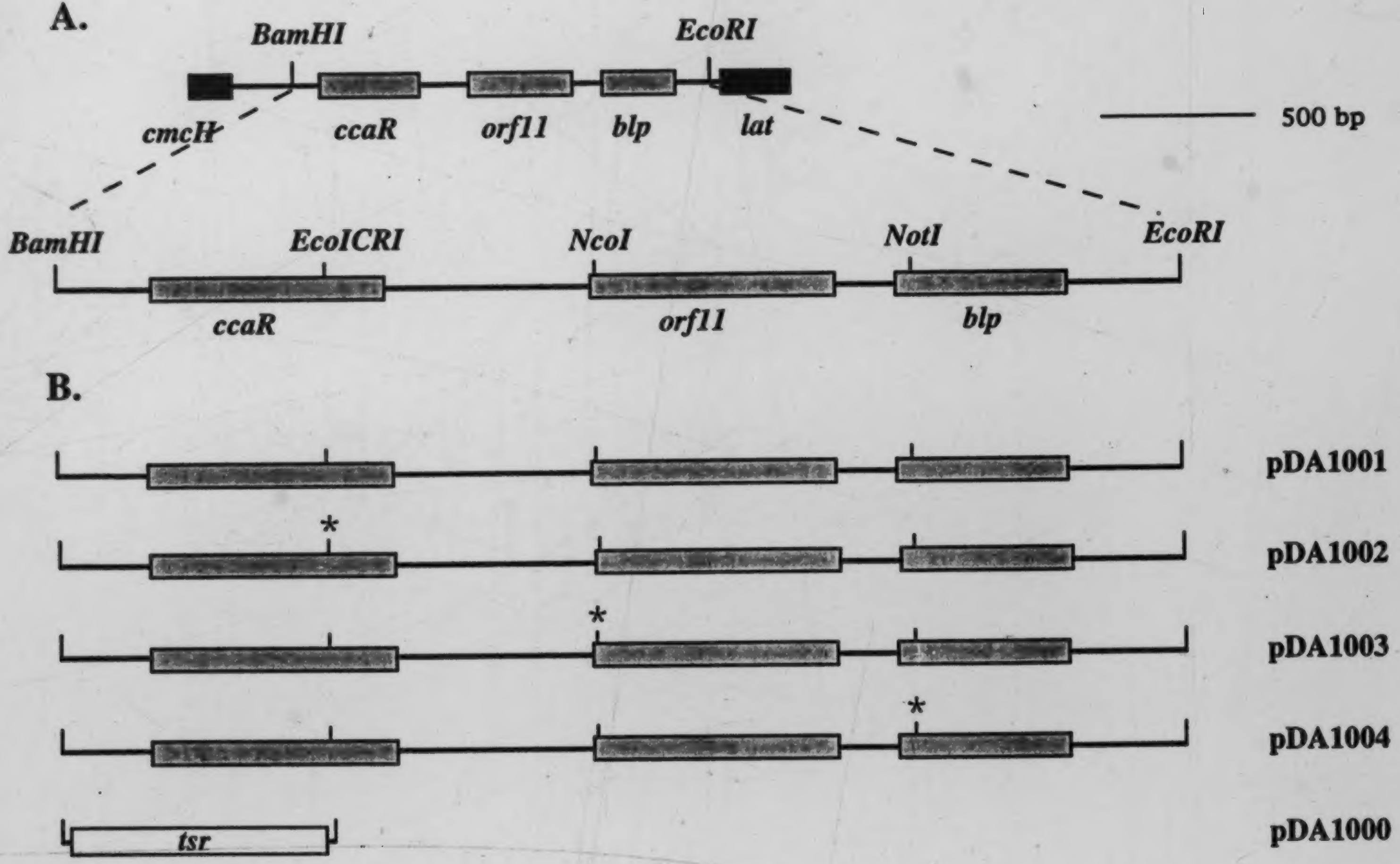
V.8.1.1 Characterization of cephamycin C production by complemented Δ*cob*::*tsr*

Two independent strains generated from each transformation of the Δ*cob*::*tsr*A1 mutant were grown in TSBS. Samples were taken after 48 hours and analyzed for the production of cephamycin.

Complementation with the wildtype *cob* genes in the plasmid pDA1001 rescued cephamycin production by the Δ*cob*::*tsr*A1 mutant but the pDA1002 plasmid carrying the *ccaR* mutant fragment or the negative control plasmid pDA1000 were not able to complement the Δ*cob*::*tsr*A1 mutant. Cephamycin production was only restored to 29% of wildtype levels when transformed with the pDA1001 plasmid. The constructs carrying mutations in the *orf11* gene (pDA1003) or *blp* gene (pDA1004) were also able to return cephamycin production but again only to 23% and 16% of wildtype levels, respectively. The level of complementation by the mutated *orf11* gene (pDA1003) and *blp* gene (pDA1004) was very similar to the complementation levels achieved by the wildtype *cob* genes (pDA1001).

Figure 45. Map of the *cob* gene-containing DNA and pSET152-based recombinant plasmids containing the *cob* genes. **A.** Restriction map of *ccaR*, *orf11* and *blp* containing region and unique sites present within this region. **B.** Diagram of wildtype and mutant constructs created in pSET152. The asterisks represent the location of the stop codons generated by the translational stop fragment inserted into these genes.

The pDA1001 vector contains the wildtype *ccaR*, *orf11* and *blp* genes, the pDA1002 vector contains wildtype *orf11* and *blp* genes but a mutated *ccaR* gene, pDA1003 vector contains wildtype *ccaR* and *blp* genes but a mutated *orf11* gene and the pDA1004 vector contains wildtype *ccaR* and *orf11* genes but a mutated *blp* gene. The negative control plasmid pDA1000 was also included in this figure.


A.
BamHI
EcoRI
cmcH
ccaR
orf11
blp
lat
500 bp
BamHI
EcoICRI
NcoI
NotI
EcoRI
ccaR
orf11
blp
B.
*
*
*
pDA1001
pDA1002
pDA1003
pDA1004
tsr
pDA1000

V.8.1.2 Characterization of production of other β-lactam antibiotics by complemented Δ*cob*::*tsr*

Two independent strains generated from each transformation of the Δ*cob*::*tsr*A1 mutant and the wildtype *S. clavuligerus* control were grown in TSBS, SA and SOY media. Samples were taken at 72 hours and analyzed for the production of cephamycin (all three media), clavulanic acid (SA and SOY media) or clavams (SOY medium only). The results of the β-lactam antibiotic production were summarized in Table 17.

Transformation with the pDA1002 plasmid containing the mutant *ccaR* fragment and the negative control plasmid pDA1000 were unable to return production of cephamycin, clavulanic acid or clavams. However, when the Δ*cob*::*tsr*A1 mutant was complemented with constructs containing a functional *ccaR* gene (pDA1001, pDA1003 and pDA1004) all returned production of cephamycin in all media and also returned production of clavulanic acid and clavams to the Δ*cob*::*tsr*A1 mutant.

V.8.2 Complementation of Δ*ccaR*::*tsr* mutant with the *ccaR* gene

The *ccaR* mutant sets characterized by western blot analysis behaved as regulatory mutants lacking not only the CcaR protein but also the cephamycin biosynthetic proteins and as a result were unable to produce cephamycin. Complementation of the Δ*cob*::*tsr*A1 mutant demonstrated that only the return of a functional *ccaR* gene was essential for the production of cephamycin and the other β-lactam antibiotics. Therefore complementation of a Δ*ccaR*::*tsr* mutant with a functional *ccaR* gene would also be expected to return cephamycin production and production of the biosynthetic enzymes.

To test this hypothesis the Δ*ccaR*::*tsr*A1 mutant was transformed with pSET152-based vectors (Fig. 46B) containing the wildtype *ccaR* gene fragment (pDA1006), two different *ccaR* mutant fragments with the TSF in *Eco*47III near 5' end of the ORF (pDA1007) or the TSF in *Eco*ICRI near 3' end of the ORF (pDA1008) or the negative control plasmid containing the *tsr* marker (pDA1000). Introduction of the TSF into the *Eco*47III and the *Eco*ICRI sites of the *ccaR* gene used in complementation studies would be expected to create truncated CcaR proteins, 38 (4.2 KD) and 191 (21.1 KD) amino acids in length respectively, instead of forming a full length CcaR protein of 256 amino acids.

V.8.2.1 Characterization of cephamycin C production by the complemented Δ*ccaR*::*tsr* mutant

Two independent strains generated from each transformation of the Δ*ccaR*::*tsr*A1 mutant were grown in TSBS. Samples were taken after 48 hours and analyzed for the

Table 17. Production of β-lactam antibiotics by complemented Δ*cob::tsr* mutants

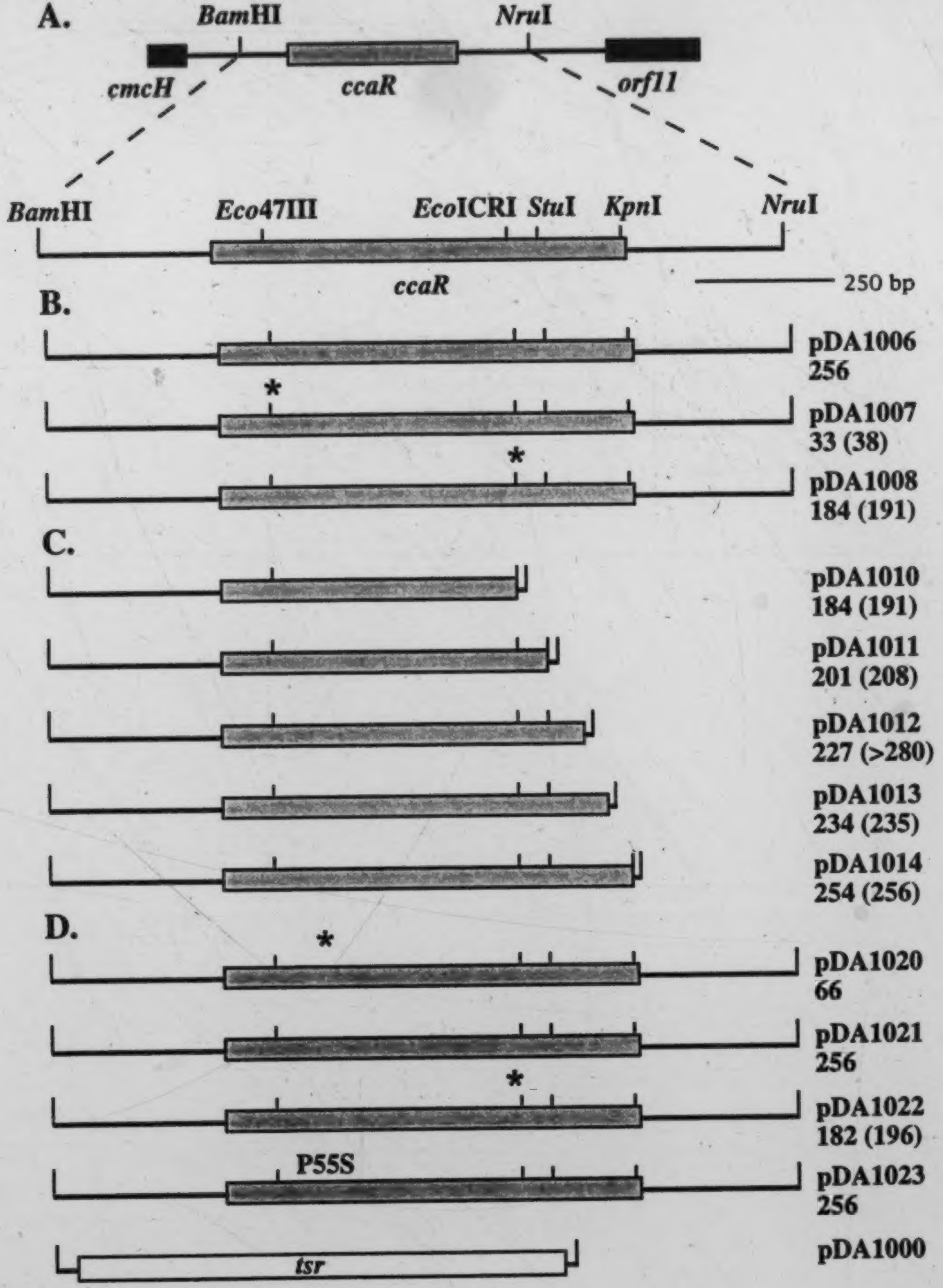
	TSBS	Starch Asparagine		SOY		
Strain	Cephamycin C	Cephamycin C	Clavulanic acid	Cephamycin C	Clavulanic acid	Clavams[a]
wildtype	+	+	+	+	+	+
Δ*cob*::*tsr*A1 pDA1000A	-	-	-	-	-	-
Δ*cob*::*tsr*A1 pDA1000B	-	-	-	-	-	-
Δ*cob*::*tsr*A1 pDA1001A	+	+	+	+	+	+
Δ*cob*::*tsr*A1 pDA1001B	+	+	+	+	+	+
Δ*cob*::*tsr*A1 pDA1002A	-	-	-	-	-	-
Δ*cob*::*tsr*A1 pDA1002B	-	-	-	-	-	-
Δ*cob*::*tsr*A1 pDA1003A	+	+	+	+	+	+
Δ*cob*::*tsr*A1 pDA1003B	+	+	+	+	+	+
Δ*cob*::*tsr*A1 pDA1004A	+	+	+	+	+	+
Δ*cob*::*tsr*A1 pDA1004B	+	+	+	+	+	+

[a] Clavams includes both carboxy-2-clavam and hydroxymethyl clavam and both metabolites are detected by HPLC analysis.

Figure 46. Map of the *ccaR* gene-containing DNA and pSET152-based recombinant plasmids containing the *ccaR* gene. **A.** Restriction map of *ccaR* containing region and unique sites present within this region. **B.** Diagram of wildtype and *ccaR* mutant constructs inserted into pSET152. The asterisks represent the location of the stop codons generated by the translational stop fragment inserted into these genes. The pDA1006 vector contains the wildtype *ccaR* gene, the pDA1007 vector contains the *ccaR* gene mutated by the insertion translational stop fragment in the 5' end of the gene and the pDA1008 vector contains the *ccaR* gene mutated by the insertion translational stop fragment in the 3' end of the gene.

The number under the names of each pSET152-based construct represents the size of the CcaR protein in amino acids while the number in brackets is the total size of the hybrid-protein which includes the non-CcaR amino acids that are coded by the translational stop fragment or the multiple cloning site sequences.

C. Diagram of the truncated *ccaR* genes created by exoIII deletions or restriction enzyme digestion. The pDA1010 vector contains the *ccaR* gene truncated at the *Eco*ICRI site, the pDA1011 vector contains the *ccaR* gene truncated at the *Stu*I site, pDA1012 vector contains the *ccaR* gene truncated by exoIII deletion, pDA1013 vector also contains the *ccaR* gene truncated by exoIII deletion and the pDA1014 vector contains the *ccaR* gene truncated at the *Kpn*I site. **D.** Diagram of *ccaR* alleles cloned from the genomic DNA of the random mutagenesis mutants and inserted into pSET152. The asterisks represent the location of the natural stop codons present in the *ccaR* alleles while P55S represents the proline to serine change at amino acid 55. The pDA1020 vector contains the mutated *ccaR* allele isolated from the *dclX*68 mutant, the pDA1021 contains the wildtype *ccaR* allele isolated from the NTG 26 mutant, pDA1022 vector contains the mutated *ccaR* allele isolated from the NTG 32 mutant, and the pDA1023 vector contains the mutated *ccaR* allele isolated from the NTG 83 mutant.


A.
*Bam*HI
*Nru*I
cmcH
ccaR
orf11
*Bam*HI
*Eco*47III
*Eco*ICRI
*Stu*I
*Kpn*I
*Nru*I
ccaR
250 bp
B.
pDA1006
256
*
pDA1007
33 (38)
*
pDA1008
184 (191)
C.
pDA1010
184 (191)
pDA1011
201 (208)
pDA1012
227 (>280)
pDA1013
234 (235)
pDA1014
254 (256)
D.
*
pDA1020
66
pDA1021
256
*
pDA1022
182 (196)
P55S
pDA1023
256
tsr
pDA1000

production of cephamycin. Also at this time the cultures were harvested and cell-extracts were prepared for western analysis.

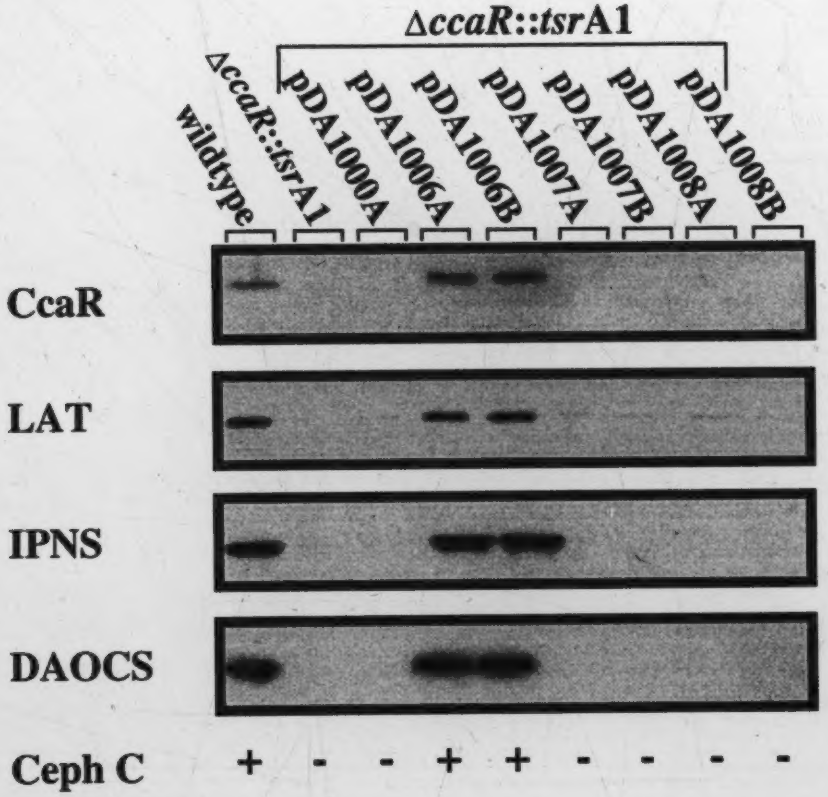
Complementation with plasmid pDA1006 containing the wildtype *ccaR* gene rescued cephamycin production to the Δ*ccaR*::*tsr*A1 mutant (Fig. 47). Cephamycin production was restored to 89% of wildtype levels. Western analysis using CcaR, LAT, IPNS and DAOCS-specific antibodies also verified that complementation returned the production of the CcaR protein and also the cephamycin biosynthetic enzymes. Strains created by transformation with plasmids pDA1007 and pDA1008 containing mutated versions of the *ccaR* gene or the negative control plasmid pDA1000 were not able to complement the Δ*ccaR*::*tsr*A1 mutant and return cephamycin production. Complementation with the mutant *ccaR* genes or the negative control plasmid also did not return production of the CcaR protein or the biosynthetic enzymes. Western analysis did not show any evidence of the truncated form of CcaR from the pDA1008 plasmid which should be about 21 KD in size and therefore instability of the truncated products may prevent complementation.

Faint bands were sometimes observed just below the LAT protein band and were believed to be due to cross-reaction of the LAT antibody but it could also suggest that the production of the LAT protein is only severely decreased but not eliminated. Therefore, the faint bands present (Fig. 47, lanes 6-9) may be the detection of a cross-reacting protein or very low levels of LAT protein production.

V.8.3 Complementation of the Δ*ccaR*::*tsr* mutant with truncated *ccaR* genes

CcaR and its homologous proteins show many conserved amino acids spread throughout the primary sequence (see Fig. 13). The uniform distribution of the conserved amino acids makes it difficult to predict the positions of the domains responsible for the function of these transcriptional activators. Only recently has it been speculated that the location of the amino acid motif responsible for CcaR binding to DNA is in the N-terminus of the protein (228). The following *ccaR* gene constructs were prepared in a preliminary attempt to determine which regions of the CcaR protein were most important for activity. It was hoped that a mutant CcaR protein capable of physically binding to the DNA but unable to activate transcription would be found. Since complementation of the Δ*ccaR*::*tsr*A1 mutants with a wildtype *ccaR* gene rescued cephamycin production and also returned production of the cephamycin biosynthetic enzymes, the function of the mutant CcaR proteins could be determined via complementation.

Figure 47. Complementation of the Δ*ccaR*::*tsr*A1 mutant with pSET152-based recombinant plasmids containing the *ccaR* gene. Western blot analysis and bioassay for production of cephamycin were used to assess complementation. The designation A or B represent two independent strains generated during the same transformation. Approximately 5 μg of cell-extract protein from each strain, harvested after 48 hours growth, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the CcaR, LAT, IPNS and DAOCS proteins. The ability of each strain to produce cephamycin C (Ceph C) was determined by bioassay and the results were scored as: + for production and - for lack of production.


ΔccaR::tsrA1
wildtype
ΔccaR::tsrA1
pDA1000A
pDA1006A
pDA1006B
pDA1007A
pDA1007B
pDA1008A
pDA1008B
CcaR
LAT
IPNS
DAOCS
Ceph C
+
-
-
+
+
-
-
-
-

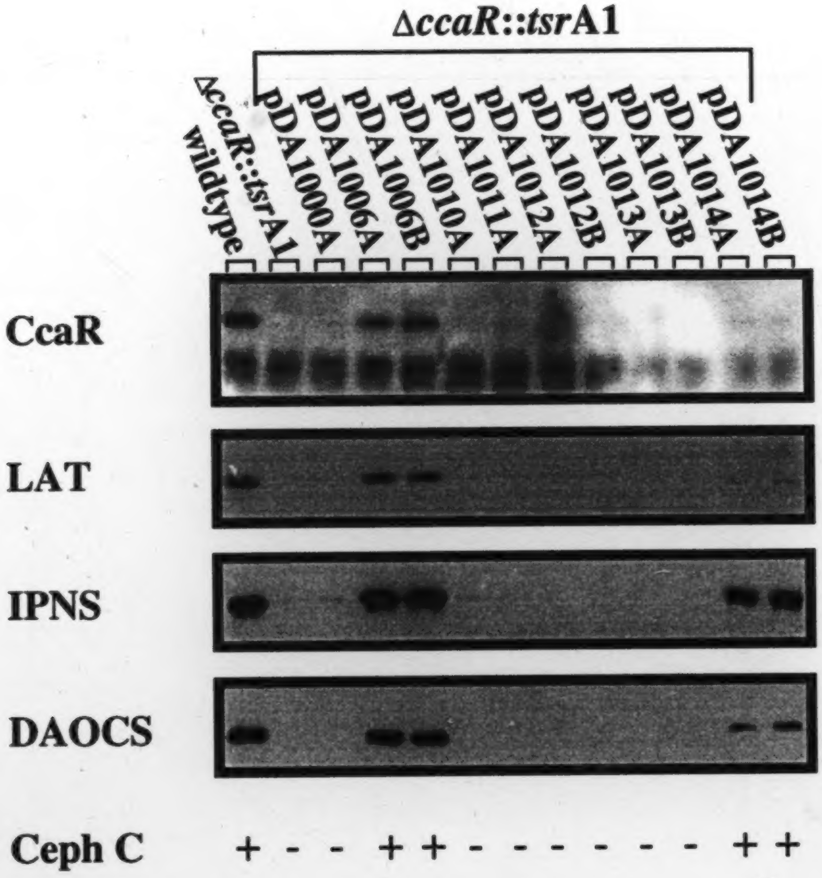
The mutant genes were created by exonuclease III digestion to generate deletions or by self-ligation of the plasmid pDA150 after digestion with a unique restriction enzyme within the gene and the plasmid MCS. These mutations would remove amino acids from the C-terminal end of the protein and by transforming these plasmids into a Δ*ccaR*::*tsr* mutant it would be possible to demonstrate the effect of the truncation on the ability of the "CcaR" protein to function as an activator. The deleted CcaR derivatives were inserted into pSET152 and the resulting recombinant plasmids were transformed into the Δ*ccaR*::*tsr*A1 mutant. The plasmid pDA1010 contained a mutant gene with the C-terminus removed from the *Eco*ICRI site, pDA1011 contained a mutant gene with the C-terminus removed from the *Stu*I site and the plasmid pDA1014 contained a mutant gene with the C-terminus removed from the *Kpn*I site (Fig. 46C). These plasmids would result in CcaR-hybrid proteins of 191 (21.1 KD), 208 (23.2 KD) and 256 amino acids (28.6 KD), respectively. The resulting proteins were hybrids because each contained C-terminal amino acids that were coded for by MCS sequence before a stop codon was reached. The plasmids pDA1013 and pDA1012 have mutant genes with the C-terminus removed by exoIII digestion from regions in between the *Stu*I site and the *Kpn*I site and produce proteins of 237 (26.2 KD) and greater than 280 amino acids (greater than 32.7 KD), respectively. The Δ*ccaR*::*tsr*A1 mutant was also transformed with the positive (pDA1006) and negative control plasmids (pDA1000) containing the wildtype *ccaR* gene and the *tsr* marker, respectively.

V.8.3.1 Characterization of cephamycin C production by the Δ*ccaR*::*tsr* mutant complemented with truncated *ccaR* genes

Two independent strains generated from each transformation of the Δ*ccaR*::*tsr*A1 mutant were grown in TSBS. Samples were taken after 48 hours and analyzed for the production of cephamycin. Also at this time the cultures were harvested and cell-extracts were prepared for western analysis.

Complementation with plasmid pDA1006 containing the wildtype *ccaR* gene and the pDA1014 plasmid was able to rescue cephamycin production in the Δ*ccaR*::*tsr*A1 mutant (Fig. 48). Cephamycin production was however only restored to 89% and 24% wildtype levels, respectively. Complementation with pDA1006 was able to return both the production of the CcaR protein and also the cephamycin biosynthetic enzymes but complementation with pDA1014 produced near undetectable amounts of the CcaR protein and greatly reduced amounts of the biosynthetic enzymes. The pDA1014 plasmid contains a *ccaR* allele missing only two CcaR amino acids which were replaced with two MCS coded amino acids at the C-terminal end. The strains created by transformation

Figure 48. Complementation of the Δ*ccaR*::*tsr*A1 mutant with pSET152-based recombinant plasmids containing truncated *ccaR* genes. Western blot analysis and bioassay for production of cephamycin were used to assess complementation. The designation A or B represent two independent strains generated during the same transformation. Approximately 5 μg of cell-extract protein from each strain, harvested after 48 hours growth, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the CcaR, LAT, IPNS and DAOCS proteins. The ability of each strain to produce cephamycin C (Ceph C) was determined by bioassay and the results were scored as: + for production and - for lack of production.


ΔccaR::tsrA1
wildtype
ΔccaR::tsrA1
pDA1000A
pDA1006A
pDA1006B
pDA1010A
pDA1011A
pDA1012A
pDA1012B
pDA1013A
pDA1013B
pDA1014A
pDA1014B
CcaR
LAT
IPNS
DAOCS
Ceph C
+ - - + + - - - - - - + +

with the plasmids pDA1010, pDA1011, pDA1012 and pDA1013 containing mutated versions of the *ccaR* gene with larger deletions of the *ccaR* ORF than pDA1014 were unable to rescue cephamycin production. Neither these plasmids nor the negative control plasmid pDA1000 were able to produce the biosynthetic enzymes or truncated versions of the CcaR protein.

V.8.4 Complementation of the Δ*ccaR::tsr* mutant with mutant *ccaR* alleles

Complementation of the Δ*ccaR*::*tsr*A1 mutants with a wildtype *ccaR* gene rescued cephamycin production and also returned production of the biosynthetic enzymes to this mutant. Transformation of the various NTG mutant *ccaR* alleles into the Δ*ccaR*::*tsr*A1 mutant would test the ability of these alleles to complement in a background which was otherwise isogenic to wildtype *S. clavuligerus*. The NTG mutants and the *dclX*68 mutant were generated by random mutagenesis and may contain multiple mutations in their genomes. The NTG 26 mutant for example is an amino acid auxotroph, is unable to sporulate well and also has the mutation(s) which affect antibiotic biosynthesis.

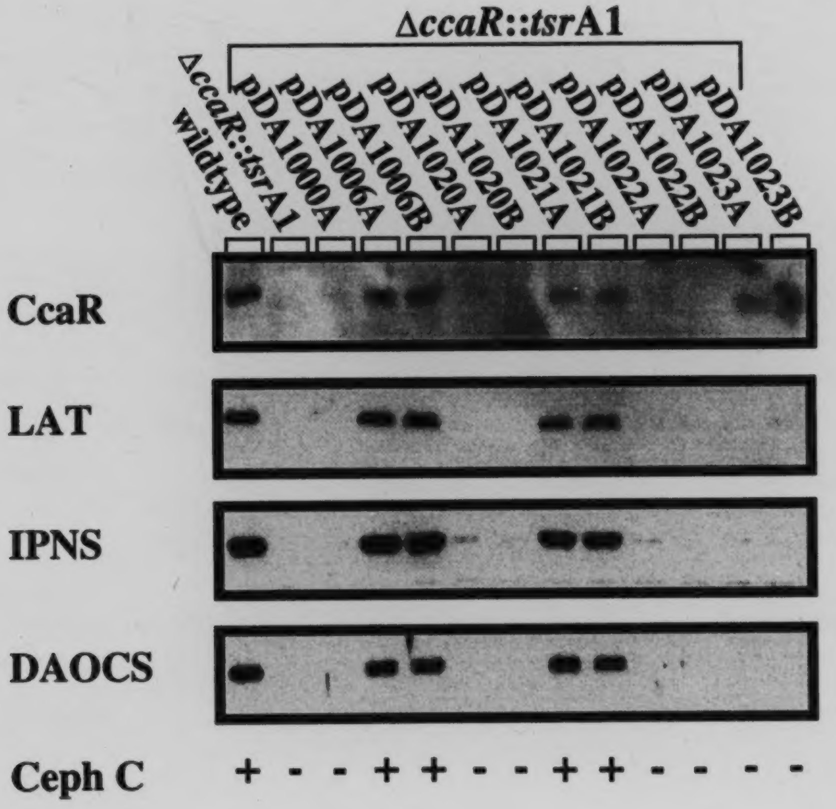
The Δ*ccaR*::*tsr*A1 mutant was transformed with pSET152-based vectors (Fig. 46D) containing the *ccaR* alleles cloned from the *dclX*68 (pDA1020), NTG 26 (pDA1021), NTG 32 (pDA1022) and NTG 83 (pDA1023) mutants as well as the positive (pDA1006) and negative control plasmids (pDA1000) containing the wildtype *ccaR* gene and the *tsr* marker, respectively.

V.8.4.1 Characterization of cephamycin C production by the Δ*ccaR::tsr* mutant complemented with mutant *ccaR* alleles

Two independent strains generated from each transformation of the Δ*ccaR*::*tsr*A1 mutant were grown in TSBS. Samples were taken after 48 hours and analyzed for the production of cephamycin. Also at this time the cultures were harvested and cell-extracts were prepared for western analysis.

Complementation with plasmid pDA1006 containing the wildtype *ccaR* gene and the pDA1021 plasmid containing the *ccaR* allele from NTG 26 rescued cephamycin production by the Δ*ccaR*::*tsr*A1 mutant (Fig. 49). Cephamycin production was however only restored to 89% and 15% of wildtype levels, respectively. Complementation was able to return both the production of the CcaR protein and also the cephamycin biosynthetic enzymes. Strains created by complementation with the plasmids pDA1020, pDA1022 and pDA1023 containing mutated versions of the *ccaR* gene from *dclX*68, NTG 32 and NTG 83 mutants, respectively or the negative control plasmid pDA1000

Figure 49. Complementation of the Δ*ccaR*::*tsr*A1 mutant with pSET152-based recombinant plasmids containing mutant *ccaR* genes. Western blot analysis and bioassay for production of cephamycin were used to assess complementation. The designation A or B represent two independent strains generated during the same transformation. Approximately 5 μg of cell-extract protein from each strain, harvested after 48 hours growth, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the CcaR, LAT, IPNS and DAOCS proteins. The ability of each strain to produce cephamycin C (Ceph C) was determined by bioassay and the results were scored as: + for production and - for lack of production.


ΔccaR::tsrA1
wildtype
ΔccaR::tsrA1
pDA1000A
pDA1006A
pDA1006B
pDA1020A
pDA1020B
pDA1021A
pDA1021B
pDA1022A
pDA1022B
pDA1023A
pDA1023B
CcaR
LAT
IPNS
DAOCS
Ceph C
+ - - + + - - + + - - - -

were not able to complement the Δ*ccaR*::*tsr*A1 mutant and return cephamycin production. Complementation with the mutant *ccaR* alleles from the *dclX*68 or NTG 32 mutants or the negative control plasmid did not return production of the CcaR protein or the biosynthetic enzymes. Complementation with the pDA1023 plasmid containing the NTG 83 mutant *ccaR* allele returned production of the CcaR protein as detected by western analysis but apparently this mutant CcaR protein was unable to activate transcription of the cephamycin biosynthetic genes and return production of cephamycin. The proline to serine mutation at amino acid 55 of the NTG 83 *ccaR* allele occurred in a proline residue which is conserved in all nine of the homologous transcriptional regulators except for the EmbR protein (see Fig. 13). As well, the amino acid residue directly adjacent to P55 was also a proline which was conserved in all of the transcriptional regulators.

V.8.5 Complementation of NTG mutants

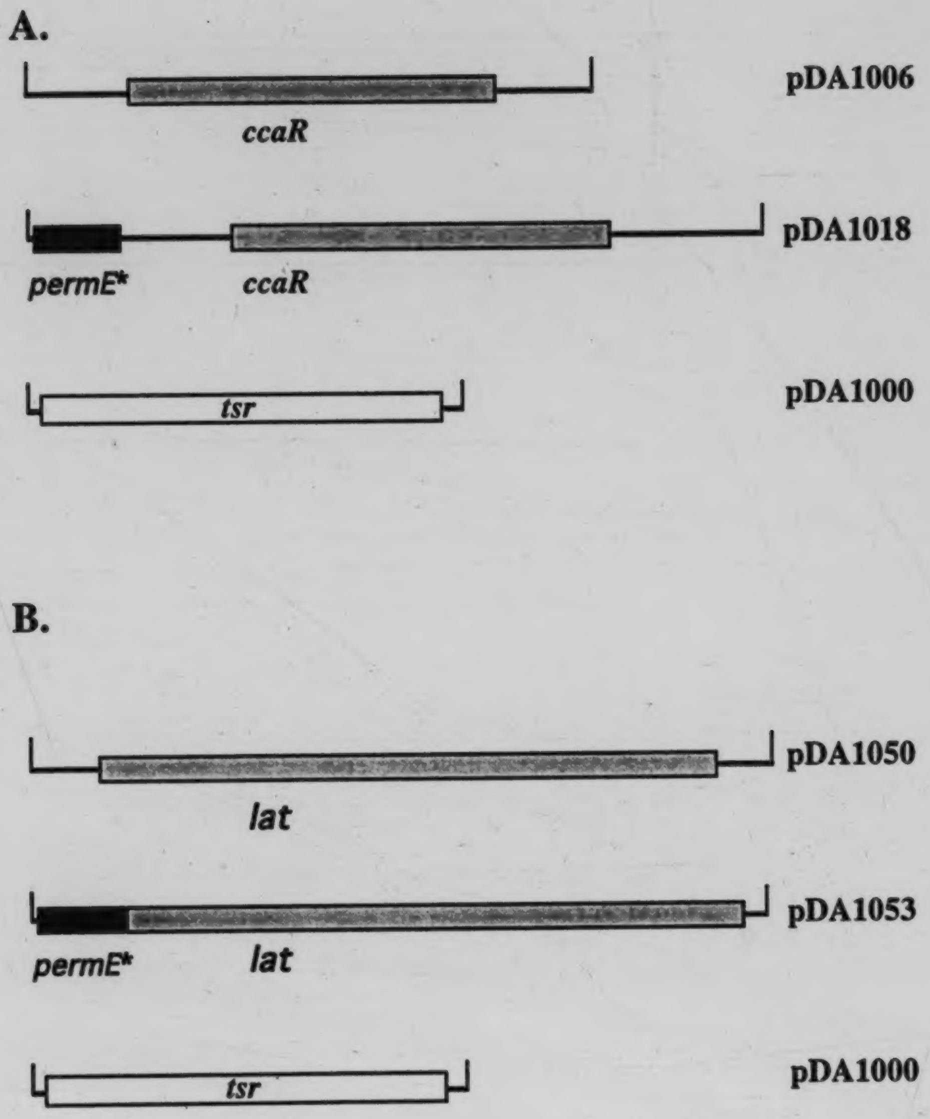
Most of the NTG mutants characterized by western blot analysis behaved as apparent regulatory mutants lacking not only the CcaR protein but also the cephamycin biosynthetic proteins and as a result were unable to produce cephamycin. When the *ccaR* alleles from the NTG 26, NTG 32 and NTG 83 mutants were sequenced the NTG 26 allele contained no mutations. As well, the NTG 83 mutant *ccaR* allele gave rise to a non-functional mutant CcaR protein in the Δ*ccaR*::*tsr* mutant when complemented with the pDA1023 plasmid but this mutant protein was not detected in the NTG 83 mutant itself (see Fig. 41).

Complementation of the Δ*ccaR*::*tsr* mutant with a functional *ccaR* gene confirmed that these mutants were capable of producing cephamycin at near wildtype levels when a single copy of the *ccaR* gene was supplied *in trans*. This approach could also be used to determine whether the NTG mutants were *ccaR* mutants that were capable of being complemented with a single copy of the *ccaR* gene. The NTG mutants were transformed with pSET152-based vectors (Fig. 50A) containing the wildtype *ccaR* gene fragment (pDA1006), the wildtype *ccaR* gene fragment with the *ermE** promoter fragment inserted in tandem with the *ccaR* promoter (pDA1030) or the negative control plasmid containing the *tsr* marker (pDA1000). The tandem promoter construct present in pDA1030 should allow expression of the CcaR protein under conditions where the *ccaR* promoter would not normally be expressed. Unfortunately, transformation of the NTG 83 strain with any of these pSET152 constructs was unsuccessful.

Attempts were made to prepare a better unregulated version of the *ccaR* gene than was present in pDA1030, by replacing the *ccaR* promoter with an expression cassette containing the strong constitutive *ermE** promoter and synthetic Shine-Dalgarno sequence

Figure 50. pSET152-based recombinant plasmids containing the *ccaR* or *lat* genes. **A.** Diagram of wildtype *ccaR* gene and the tandem promoter construction inserted into pSET152. The pDA1006 vector contains the wildtype *ccaR* gene and the pDA1018 vector contains the wildtype *ccaR* and the *ermE** promoter.

B. Diagram of the wildtype *lat* gene and the *ermE** high level expression construct inserted into pSET152. The pDA1050 vector contains the *lat* promoter and *lat* gene while the pDA1053 vector contains the *ermE** promoter fragment attached to the *lat* gene. The pDA1000 plasmid containing the *tsr* marker was included since it serves as the negative control in these studies.


A.
pDA1006
ccaR
pDA1018
permE*
ccaR
pDA1000
tsr
B.
pDA1050
lat
pDA1053
permE*
lat
pDA1000
tsr

removed from the *Streptomyces* expression vector pHM8a (153). This expression cassette was fused with the *ccaR* gene at an artificial *Nde*I site created at the ATG start codon as predicted by Walters et al. (221). When this constitutive *ccaR* gene, cloned into pSET152, was transformed into wildtype *S. clavuligerus* or the Δ*ccaR*::*tsr* mutant the resulting transformants regenerated more slowly than normal. These transformants also grew poorly as patches on MYM plates and TSBS liquid cultures. The reason for the toxicity of this construct was not attributable to cephamycin overproduction since the transformation of the cephamycin non-producing Δ*lat*::*tsr* mutant with this plasmid resulted in the same phenotype.

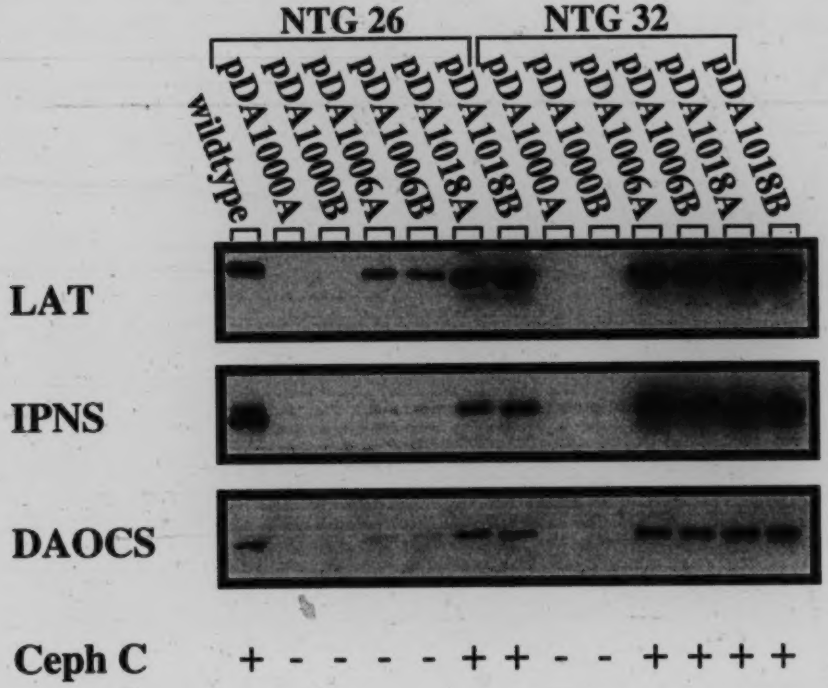
V.8.5.1 Characterization of cephamycin C production by the complemented NTG 26 and NTG 32 mutants

Two independent strains generated from each transformation of the NTG 26 or the NTG 32 mutants with the various *ccaR* expression constructs were grown in TSBS. Samples were taken after 48 hours and analyzed for the production of cephamycin. Also at this time the cultures were harvested and cell-extracts were prepared for western analysis.

Transformation with plasmid pDA1006 containing the wildtype *ccaR* gene or the negative control plasmid pDA1000 was not able to complement the NTG 26 mutant and rescue cephamycin production (Fig. 51). Although no cephamycin production was detected in the NTG 26 strains created by complementation with pDA1006 low levels of the LAT, IPNS and DAOCS cephamycin biosynthetic enzymes were detected although they were absent in strains created by transformation with pDA1000. The CcaR protein was not detected in either of these transformants (data not shown). Cephamycin production by the NTG 26 mutant was rescued by complementation with the pDA1030 plasmid containing the tandem promoter construction. Cephamycin production was however only restored to 1.5% of wildtype levels. These transformants also produce high levels of the LAT, IPNS and DAOCS cephamycin biosynthetic enzymes as well as the CcaR protein.

Transformation of the NTG 32 mutants with the plasmid pDA1006 containing the wildtype *ccaR* gene or the tandem promoter construct pDA1030 rescued cephamycin production (Fig. 51). Cephamycin production was returned to 44% and 270% of wildtype levels for pDA1006 and pDA1030 respectively. This complementation returns the production of the CcaR protein (data not shown) and also the cephamycin biosynthetic enzymes. Complementation of the NTG 32 mutant with the negative control plasmid

Figure 51. Complementation of the NTG 26 and NTG 32 mutants with pSET152-based recombinant plasmids containing the *ccaR* gene. Western blot analysis and bioassay for production of cephamycin were used to assess complementation. The designation A or B represent two independent strains generated during the same transformation. Approximately 5 µg of cell-extract protein from each strain, harvested after 48 hours growth, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the LAT, IPNS and DAOCS proteins. The ability of each strain to produce cephamycin C (Ceph C) was determined by bioassay and the results were scored as: + for production and - for lack of production.


NTG 26
NTG 32
wildtype
pDA1000A
pDA1000B
pDA1006A
pDA1006B
pDA1018A
pDA1018B
pDA1000A
pDA1000B
pDA1006A
pDA1006B
pDA1018A
pDA1018B
LAT
IPNS
DAOCS
Ceph C
+ - - - - + + - - + + + +

pDA1000 did not return production of the CcaR protein, the biosynthetic enzymes or cephamycin.

V.9 Strength of the *lat* promoter in wildtype and *ccaR* mutants

The *ccaR* mutants created by homologous recombination were unable to produce the LAT, ACVS, IPNS, IPNE or DAOCS biosynthetic enzymes and therefore also lacked cephamycin production. Complementation of the *ccaR* mutant with a functional *ccaR* gene resulted in a return of the ability to produce the cephamycin biosynthetic enzymes and cephamycin. These results along with the fact that the CcaR protein shows homology to *Streptomyces* transcriptional activators suggests that inactivation of the *ccaR* gene results in the loss of the biosynthetic enzymes due to regulation at the transcriptional level.

If the CcaR protein was responsible for the regulation of biosynthetic enzyme production and the species-specific activation of the *lat* promoter in *S. clavuligerus*, a promoter probe construct containing the *lat* promoter in a *ccaR* mutant background would be expected to display a similar lack of activation.

To test this hypothesis two *ccaR* mutant backgrounds and wildtype *S. clavuligerus* were transformed with the promoter probe plasmids pDA2000 and pDA2003 (see Fig. 22). The plasmid pDA2003 contains the *lat* promoter as well as part of the *lat* coding region which has been proposed to contain the putative *pcbAB* promoter.

Promoter probe studies relied on pDA2000-based plasmids which integrate into the genome of *S. clavuligerus*. Verification of plasmid integration was confirmed by genomic PCR (see IV.2.1). The amplified PCR product from gDNA of the transformed strain was compared to the product amplified from free plasmid isolated from *E. coli*.

V.9.1 Promoter probe analysis of the *lat* promoter

Two independent strains generated from each transformation of the wildtype and *ccaR* mutant strains were grown in TSBS. Cultures were harvested (see II.2.4.5) after 48 hours and cell-extracts were prepared for catechol 2,3 dioxygenase (C230) activity measurements (see II.2.4.5.1). Also at this time samples were taken and assayed for production of cephamycin.

The C230 specific activity of the *lat* gene containing pDA2003 plasmid in the *ccaR::apr* and the *ΔccaR::tsr* mutants was only 2.6% and 1.3% of the levels obtained from wildtype *S. clavuligerus* (Table 18). These levels are even below that for the pDA2000 plasmid, which does not contain an insert, in either background.

Therefore in *S. clavuligerus ccaR* mutants transcription from the *lat* promoter was severely decreased, suggesting very tight control.

Table 18. Promoter probe analysis in *ccaR* mutants

strain	Catechol dioxygenase specific activity [a] strain A [b]	strain B [b]	average	average [c] wildtype
wildtype/pDA2003	123	137	130	100
wildtype/pDA2000	6.1	7.3	6.7	5.2
ccaR::*apr*A1/pDA2003	1.7	1.6	1.7	1.3
Δ*ccaR*::*tsr*A1/pDA2003	4.7	2.1	3.4	2.6
Δ*ccaR*::*tsr*A1/pDA2000	6.7	6.7	6.7	5.2

[a] C23O specific activity values are given in EU x 10^{-6}/mg protein.

[b] Specific activities are an average of two seperate fermentations for each strain.

[c] Values are calculated as a percentage of the specific activity obtained for the pDA2003 plasmid in wildtype *S. clavuligerus*.

V.10 CcaR-dependent regulation of the *lat* promoter

The *ccaR* mutants were shown to be unable to produce the LAT, ACVS, IPNS, IPNE or DAOCS biosynthetic enzymes. Promoter probe plasmids containing the *lat* promoter transformed into *ccaR* mutants showed a severe reduction in C23O activity when compared to the same plasmid construct transformed into wildtype *S. clavuligerus*. These results suggest that inactivation of the *ccaR* gene regulates the *lat* gene as well as the other CcaR-dependent cephamycin biosynthetic genes at the level of transcription.

To confirm the mechanism of CcaR-dependent regulation a second copy of the *lat* gene was added *in trans* to a Δ*ccaR*::*tsr* mutant. In the construct pDA1053, the *lat* promoter was replaced by the *ermE**-based expression cassette from the *Streptomyces* expression vector pHM8a (153). This made it possible to determine if production of LAT was CcaR-dependent in the absence of the *lat* promoter. The *ccaR*::*tsr*A1 mutant was transformed with pSET152-based vectors containing the wildtype *lat* gene (pDA1050), the *lat* gene with its native promoter replaced by the expression cassette (pDA1053) or the negative control plasmid containing the *tsr* marker (pDA1000) (Fig. 50B).

V.10.1 Characterization of LAT production by the complemented Δ*ccaR*::*tsr* mutant

Two independent mutants from each transformation of the Δ*ccaR*::*tsr*A1 mutant with the pSET152-based constructs were grown in TSBS. Samples were taken after 48 hours and analyzed for the production of cephamycin. Also at this time the cultures were harvested and cell-extracts were prepared for western analysis.

Transformants containing the plasmid pDA1050 or the negative control plasmid pDA1000 were not able to produce cephamycin. Western analysis of these complemented strains also demonstrated that no LAT or IPNS protein was produced (data not shown). Transformation of the Δ*ccaR*::*tsr*A1 mutant with the plasmid pDA1053 did not allow the mutant strain to produce cephamycin but changing the promoter allowed the LAT protein to be expressed in a CcaR-independent manner.

V.11 Time course of production of CcaR during growth of *S. clavuligerus*

If the CcaR protein is the transcriptional activator responsible for transcription of the cephamycin biosynthetic genes then production of the CcaR protein should precede the expression of the cephamycin biosynthetic enzymes it controls. Western analysis of cell-extracts from young wildtype *S. clavuligerus* cultures should show the expression of the putative transcriptional activator followed by the biosynthetic enzymes.

Spores of wildtype *S. clavuligerus* were inoculated directly into 25 ml amounts of TSBS medium. Multiple flasks were harvested at early time points to give enough material for analysis.

V.11.1 Analysis of the time course of CcaR and cephamycin biosynthetic enzyme production by western blot analysis

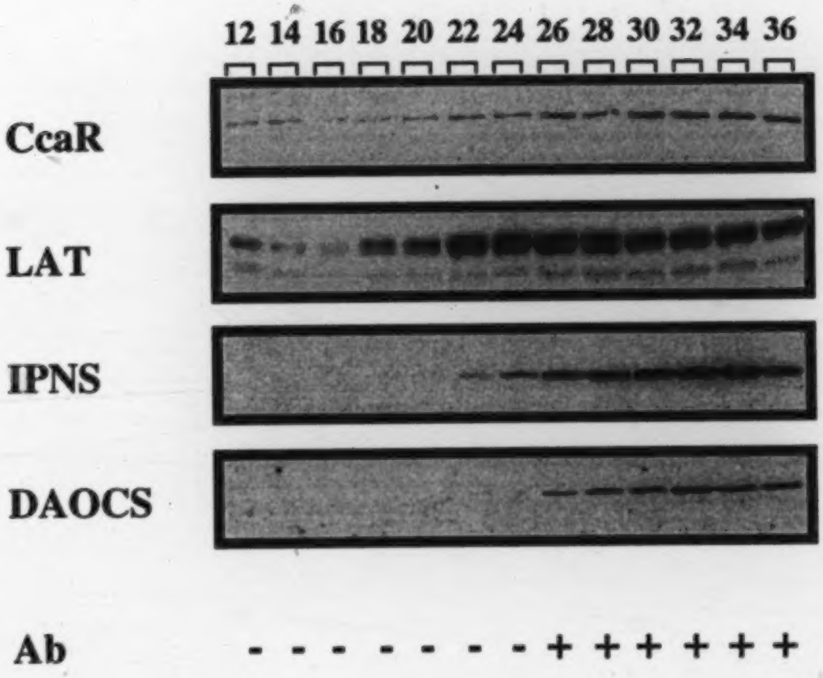
Expression of the CcaR protein preceded expression of the cephamycin biosynthetic enzymes. The appearance of the LAT and IPNS proteins was delayed relative to the CcaR protein which was present at 12 hours after inoculation (Fig. 53). There was a faint band in all of the LAT lanes which makes it appear that the LAT protein was also present at 12 hours but it actually was not present until 18 hours and the IPNS protein appeared after 22 hours. The appearance of the DAOCS protein occurred after 26 hours which was well after the early LAT and IPNS enzymes which are involved in the first steps of cephamycin biosynthesis. The appearance of penicillin/cephamycin antibiotic in the supernatant first occurred at 26 hours. Many intermediates of the cephamycin biosynthetic pathway are bioactive against the *E. coli* indicator strain so at these early time points it is unlikely that only cephamycin C is being measured.

V.12 Attempts to titrate the regulator of antibiotic biosynthesis

If the CcaR protein is responsible for the activation of transcription from the *lat* promoter the addition of multiple copies of the *lat* promoter should titrate out the transcriptional activator and create a situation where the chromosomal *lat* promoter responsible for the transcription of the *lat/pcbAB/pcbC* polycistronic mRNA would be less well transcribed and as a result less cephamycin would be made. This decreased transcription should also occur for any gene whose expression is CcaR-dependent.

To test this hypothesis wildtype *S. clavuligerus* was transformed with pUWL-KS-based plasmids containing: the *lat* promoter fragment (pDA601), the wildtype *lat* gene (pDA602) and the negative control plasmid pUWL-KS. The pUWL-KS plasmid is an *E. coli/Streptomyces* shuttle plasmid using the high copy number pIJ101 replicon for replication in *Streptomyces*. Other pIJ101 plasmids, such as pIJ486 have an estimated copy number of 40 to 300 in *S. lividans* (97), and therefore it is likely that pUWL-KS will be present in *S. clavuligerus* at a very high level.

Figure 52. Western blot analysis of an *S. clavuligerus* time course. Approximately 5 µg of cell-extract protein from each strain, harvested at each designated time, was separated by 10% SDS-PAGE and transferred to PVDF membranes. The resulting western blots were developed with polyclonal antibodies specific for the CcaR, LAT, IPNS and DAOCS proteins. The production of penicillin/cephamycin at each time point was judged by bioassay and the results were scored as: + for production and - for lack of production.


12 14 16 18 20 22 24 26 28 30 32 34 36
CcaR
LAT
IPNS
DAOCS
Ab
- - - - - - - + + + + + +

V.12.1 Characterization of cephamycin C production by *S. clavuligerus* containing high copy number plasmids

Two independent strains generated from each transformation of the wildtype *S. clavuligerus* were grown in TSBS. Samples were taken after 48 and 72 hours growth and analyzed for the production of cephamycin.

The plasmid pDA601 containing the *lat* promoter produced an average of 109% and 112% of cephamycin levels produced by wildtype *S. clavuligerus* transformed with pUWL-KS, after 48 and 72 hours growth respectively. The plasmid pDA602 containing the wildtype *lat* gene produced an average of 113% and 97% of cephamycin levels produced by wildtype *S. clavuligerus* transformed with pUWL-KS, after 48 and 72 hours growth respectively. Further analysis of the effect of these constructs on the cephamycin biosynthetic enzymes was not carried out but the wildtype *lat* gene present at high copy supported high levels of LAT protein production without any deleterious effects on cephamycin biosynthesis (data not shown).

VI. Discussion

VI.I Sequencing of the cephamycin biosynthetic cluster

The genes responsible for antibiotic biosynthesis, transport, resistance and regulation are organized into clusters in *Streptomyces* (146). The overall objective of this study was to understand more about how the production of β-lactam metabolites is regulated in *Streptomyces clavuligerus*. The complete sequence of the *Nocardia lactamdurans* cephamycin cluster appeared to be devoid of any gene directly responsible for the regulation of cephamycin C biosynthesis. A large scale sequencing project of the cephamycin cluster from *S. clavuligerus* had not been initiated and the majority of the previously sequenced genes coded for biosynthetic enzymes but several unsequenced regions remained. The only gene sequenced from the cephamycin cluster that was not involved in biosynthesis was the *pcbR* resistance gene (166). The cephamycin gene cluster and the clavulanic acid gene cluster are organized into a "super-cluster" in *S. clavuligerus* (222) and the *pcbR* gene was discovered during the sequencing of the clavulanic acid cluster (166). The cephamycin cluster sequencing component of this thesis was originally undertaken to search for potential regulators responsible for the regulation of cephamycin biosynthesis in *S. clavuligerus*.

The nucleotide sequences for several previously unsequenced regions of the cephamycin gene cluster were determined and with this new information the sequence of the entire cephamycin gene cluster was completed except for an internal region of the *pcbAB* gene. At present only 3.2 kb of the total 11 kb representing the *pcbAB* gene has been sequenced. Since the entire *pcbAB* gene had been sequenced from *N. lactamdurans* (52) as well as from other penicillin/cephalosporin producing organisms the sequence of the *S. clavuligerus* gene was not determined.

Seven open reading frames (ORFs) not previously described in *S. clavuligerus* were discovered as a result of the sequencing project and five additional undescribed ORFs were discovered but their sequences were published by others during or after completion of the sequencing project. The new ORFs discovered within the newly determined sequence were detected by codon preference analysis (25, 60) and database homology searches (9, 10). The majority of the new genes are deduced from this DNA sequence and are presumed to produce their respective proteins. The new ORFs were labeled with the gene designations that were felt to be appropriate based upon homology to known genes from *Streptomyces* or other organisms.

The total amount of compiled sequence generated from this study was greater than 26 kb and when the unsequenced region of *pcbAB* gene as well as the sequences of the *pcbC* and *pcbR* genes were included it was estimated that the entire *S. clavuligerus* gene cluster would be approximately 38 kb. This claim may be an overestimation as some uncertainty remains as to whether *sclU* and/or *rhsA* represent the end of the gene cluster or whether one or both of these genes should be considered part of the biosynthetic cluster. The *sclU* gene encodes a protein containing a number of transmembrane domains capable of inserting the 27 KD protein into the cytoplasmic membrane. A role for the SclU protein in some aspect of cephamycin biosynthesis is uncertain since the SclU protein does not show significant homology to any known proteins during database searches. Until mutants of the *sclU* gene have been generated by gene replacement it is very hard to estimate whether the *sclU* gene has a role in β-lactam biosynthesis and should be considered as part of the biosynthetic cluster.

The *rhsA* gene from *S. clavuligerus* has only been partially sequenced but the C-terminal end of the putative RhsA protein has strong homology to the Rhs proteins from *Escherichia coli*. The function of the *E. coli* Rhs proteins is unknown but they have been suggested to be cell surface ligand-binding proteins (94). The regions are also known to be recombination hot spots and the Rhs elements from *E. coli* also contain distinct components with %GC compositions which are very different from that of the surrounding genome which suggests that they recently originated from outside of *E. coli*. The presence of a recombination hot spot at one potential end of the cephamycin biosynthetic cluster provides speculation towards a mechanism by which the gene cluster has moved by horizontal transfer from *Streptomyces* spp. to *Streptomyces* spp.

The function of the *rhsA* gene in *S. clavuligerus* is unknown but it seems unlikely that it would be involved in cephamycin biosynthesis. In fact, mutants of the *rhsA* gene have been created in wildtype *S. clavuligerus* by homologous recombination and the resulting *rhsA* mutants have undergone preliminary characterization and they were not impaired in cephamycin biosynthesis (8). That the *rhsA* mutants were fully viable and did not display any abnormal phenotypes suggests that the gene was not essential to *S. clavuligerus* which was not unexpected since certain strains of *E. coli* do not have any of the *rhs* alleles whereas other strains do (94). The *rhsA* mutants have yet to be characterized for the ability to produce the other β-lactam metabolites but it seems unlikely that *rhsA* should be considered as part of the cephamycin biosynthetic cluster.

New genes identified as a result of these sequence analyses of gaps within the cephamycin cluster appear to encode proteins responsible for resistance, transport, biosynthesis and regulation functions in cephamycin biosynthesis. The PbpA protein likely acts as a membrane anchored resistance protein which along with the *bla* (173) and *pcbR* (166) gene products serves to protect the cell from the potentially lethal effects of the metabolites that it produces. Resistance genes also flank the cephamycin cluster in *N. lactamdurans* (51) but in *S. clavuligerus* the gene organization differs because *pcbR* has no counterpart in *N. lactamdurans*. Unexpectedly, the Pbp from the *N. lactamdurans* cluster was not identified as homologous to PbpA in the database search as the two proteins share only 19% identical amino acids.

Mutants of the *pbpA* gene have been created in both wildtype *S. clavuligerus* and the cephamycin non-producer NTG 19 mutant by gene replacement (8) whereas mutants of *pcbR* could only be generated in the NTG 19 mutant (166). Gene replacement mutants of *pbpA* displayed a hypersensitivity to benzylpenicillin and cephalosporins (8), similar to the *pcbR* mutants. This confirms that the PbpA protein confers resistance to β-lactam metabolites in *S. clavuligerus* but PbpA is apparently not required to protect the cell from the antibiotics that it produces. How the *pbpA* mutants survive under wildtype antibiotic producing conditions is unknown but these *pbpA* mutants will certainly require further characterization.

The *cmcT* gene encodes a membrane associated protein that contains a number of transmembrane domains. The CmcT protein shows homology to a number of *Streptomyces* transport proteins from the drug resistance translocator family involved in the export of antibiotics and/or resistance to antibiotics. Therefore the CmcT protein, which has a counterpart in the *N. lactamdurans* cluster, likely functions as a transporter that exports cephamycin from the cell.

The *pcd* gene represents one of the more interesting genes present in the newly described sequences. The PCD protein shows strong homology to semialdehyde dehydrogenases when analyzed by database searches. Alignment of the active site regions of the PCD protein and some of the most homologous proteins showed many amino acids which were conserved and all of the proteins contained the very important glutamic acid and cysteine active site residues. As well, many other pockets of highly conserved amino acids existed throughout the homologous proteins. Why most of the homologous proteins were from eukaryotic backgrounds was unknown and unfortunately these proteins have not been biochemically characterized to determine what substrates they recognize.

Production of the α–aminoadipate (αAA) precursors required for cephamycin biosynthesis is a two step process (119). The first step is catalyzed by lysine-ε-aminotransferase (LAT) which removes the epsilon amino group from lysine to create the intermediate 1-piperideine-6-carboxylate (cyclized form of aminoadipic semialdehyde). The second step converting 1-piperideine-6-carboxylate to αAA is catalyzed by a piperideine-6-carboxylate dehydrogenase (PCD) which has recently been purified and characterized from *S. clavuligerus* (55). However, a number of gene clusters from β-lactam producing species have now been partially or fully sequenced and no candidate dehydrogenase genes have been identified (170) which led to speculation that the reaction might be catalyzed by a primary metabolic enzyme. *pcd* may represent that missing gene, a speculation that is consistent with the location of *pcd* within the cephamycin gene cluster and with the estimated size of the purified PCD enzyme at 52.6 KD (55) which corresponds well with the deduced molecular weight of the *pcd* gene product at 54.0 KD. However, it seems unusual that since LAT and PCD are proposed to catalyze sequential steps in the pathway, their corresponding genes would not be more closely grouped, especially in view of the multicistronic transcriptional relationship that exists between *lat* and the *pcbAB* and *pcbC* genes which encode the enzymes for the next two steps in the biosynthetic pathway.

Instead, *pcd* is grouped together with genes encoding enzymes from the middle part of the pathway and is likely co-transcribed with them as a polycistronic transcript. The *pcd* gene along with *cmcT* and possibly *pbpA* genes may be transcribed as a polycistronic transcript together with the *cefD* and *cefE* genes, from a promoter located upstream of *cefD*. Kovacevic et al. (123) reported a *cefD*/*cefE* hybridizing mRNA of approximately 10 kb, which would be large enough to include all five of the genes within one transcript. Both the *bla* gene downstream of *pbpA*, and the *cmcI* gene upstream of *cefD* are oriented in the opposite direction to the group of five genes and so therefore cannot be part of the multicistronic transcript. It also seems unlikely that the *pbpA* gene would be included in the polycistronic transcript since there is greater than 1.2 kb of intergenic DNA between *pbpA* and *cmcT*. The *cmcT* and *pcd* coding regions are more likely to be part of the polycistronic transcript since the intergenic regions between the *cmcT* and *pcd* genes and the *pcd* and *cefE* genes are 429 and 15 bp, respectively. S1 nuclease mapping and/or northern analysis would be required to determine which of the genes are present on the polycistronic transcript originating from the *cefD* promoter.

A second area of concern is that no *pcd* gene was identified in the *N. lactamdurans* cephamycin cluster (consistent with the lack of cross hybridization between *pcd* and *N. lactamdurans* genomic DNA) but PCD activity was detected in *N. lactamdurans* (55). The fact that *N. lactamdurans* can form αAA despite the absence of the *pcd* gene implies that another route must be available. If that alternative route also exists in *S. clavuligerus*, then it follows that *pcd* may not actually be involved in αAA formation, but rather have some other unrecognized function which has no equivalent in *N. lactamdurans*. However, *pcd* was shown by Southern hybridization to have an equivalent in all other β-lactam species examined and none of the non-producing species, so whatever role the *pcd* gene plays it seems closely associated with the production of cephamycin.

Preliminary characterization of *pcd* mutants created by gene replacement suggested that cephamycin production was reduced to 20% of wildtype production levels but was not eliminated (8). This suggests that the *pcd* gene does encode the PCD protein required in the conversion of lysine to αAA but that another enzyme present in *S. clavuligerus* is also capable of catalyzing this reaction. This enzyme may be solely responsible for the conversion of 1-piperideine-6-carboxylate to αAA in *N. lactamdurans*.

The *cmcH*, *cmcI* and *cmcJ* genes were expected to be located within the unsequenced gaps in the *S. clavuligerus* cluster, since similar genes had previously been located in *N. lactamdurans* and are responsible for producing the late biosynthetic enzymes. The CmcH protein from *N. lactamdurans* is believed to be a 3'-hydroxymethylcephem *O*-carbamoyltransferase (54). In *N. lactamdurans* the CmcI and CmcJ proteins combine to form the CmcI:CmcJ complex and function as a hydroxylase and a methyltransferase respectively, converting *O*-carbamoyl-desacetylcephalosporin C (OCDAC) to 7-α-hydroxy-OCDAC (70) and it is likely that these three proteins have the same function in *S. clavuligerus*.

The *cmcI* and *cmcJ* genes along with the *cefF* and *cmcH* genes appear to have moved as a 'late gene cassette' from their location between the *cefD* and *ccaR* genes in *S. clavuligerus* to reside between *pcbC* and *bla* in *N. lactamdurans*. The entire group of late genes has maintained the same position and orientation which suggests that it may function as operon. This appears likely as the *cmcI*, *cmcJ*, *cefF* and *cmcH* genes from *S. clavuligerus* are separated by 48, 96 and 75 bp of intergenic sequence but again this will require mRNA analysis to confirm. As well, promoter probe analysis of the *N. lactamdurans cmcH* gene intergenic region suggested the region did not contain a functional promoter and was likely present in a polycistronic transcript (43).

Therefore *cmcH*, *cmcI* and *cmcJ* are newly identified biosynthetic genes from *S. clavuligerus* which act together with *lat*, *pcbAB*, *pcbC*, *cefD*, *cefE* and *cefF* genes (and very likely *pcd*) to make cephamycin. The *pbpA*, *cmcT* and *ccaR* genes are newly identified genes from *S. clavuligerus* that are involved in the resistance, transport and regulation of cephamycin biosynthesis, respectively. No roles for the *rhsA*, *sclU*, *orf11* or *blp* genes were apparent from the database homology searches.

VI.2 Analysis of regulatory genes in the cephamycin cluster

The sequencing project was originally undertaken to search for potential transcriptional activators for β-lactam biosynthesis. Therefore the *ccaR* gene became the focus of subsequent studies. The sequence for the transcriptional activator *dclX/ccaR* gene was determined by both Walters et al. (221) and Perez-Llarena et al. (174). Function of the *ccaR* gene was deduced from its homology to *Streptomyces* transcriptional activators but unexpectedly the most similar proteins detected in the database searches were from *Mycobacterium*. The CcaR protein shows homology to the *Streptomyces* transcriptional activators DnrI [daunorubicin (DNR); 206], ActII-ORF4 [actinorhodin; (ACT) 74], RedD [undecylprodigiosin (RED); 155], SnoA (nogalamycin; 232) and AfsR (ACT and RED; 77) that are responsible for the regulation of antibiotic biosynthesis. Functions for the *Mycobacterium* homologs are unknown. Alignment of the deduced amino acid sequences of the CcaR homologs showed that many amino acids were conserved throughout the entire length of the proteins.

As well, like many of the transcriptional activators from *Streptomyces* spp. the CcaR protein has the rare leucine-encoding TTA codon in the N-terminal end of the protein. Regulation of both antibiotic biosynthesis and sporulation via the TTA codon has been demonstrated in *S. coelicolor* (136). Interestingly, the leucine residue encoded by the rare TTA codon of the CcaR protein was conserved throughout the CcaR homologs.

The sequences of Walters et al. (*dclX*; 221) and Perez-Llarena et al. (*ccaR*; 174) are identical within the coding region but they have chosen different start codons for the gene. Since neither potential start codon had a particularly strong Shine-Dalgarno (S/D) sequence preceding it our estimation was based upon the alignment of the N-terminal ends of the homologous proteins. Since the N-termini of the DnrI, ActII-ORF4 and Y164 proteins all align directly with the CcaR GTG start codon, it was the more likely start codon. N-terminal sequencing of the purified protein would be required to give a definitive answer but none of the experiments done in this study should be affected by which is the authentic CcaR start codon.

Until recently, it was unclear how the majority of these CcaR homologous transcriptional activators would interact with DNA since they did not contain a recognizable helix-turn-helix (HTH) motif that mediates site-specific binding of proteins to DNA (150). Wietzorrek and Bibb (228) suggested that these *Streptomyces* activators along with the *Mycobacterium* EmbR protein constitute a novel family of proteins that contain an OmpR-like DNA-binding fold. The N-terminus of these proteins contain a region which resembles the DNA-binding domain of OmpR in both primary amino acid sequence and predicted secondary structure and likely functions in an analogous fashion to a HTH motif. It was proposed that these activator proteins would interact with the promoters of the regulated genes via this DNA-binding fold.

The presence of *ccaR* cross-hybridizing DNA in *Streptomyces* β-lactam antibiotic producing species but not in *N. lactamdurans* suggests that β-lactam production in *Streptomyces* spp. proceeds through a method of transcriptional regulation which is conserved throughout the genus, but not in *N. lactamdurans*. As an alternative to regulation by CcaR, regulation of cephamycin production in *N. lactamdurans* has been suggested to involve the *orf12* gene product (51) or the *bla* gene product (127) rather than *ccaR*.

Involvement of CcaR in transcriptional regulation of the cephamycin biosynthetic pathway was predicted by homology to other pathway-specific transcriptional activators from *Streptomyces* spp. and from its location within the center of the cephamycin biosynthetic cluster. An essential role for the CcaR protein in the production of both cephamycin and clavulanic acid in *S. clavuligerus* has been demonstrated (174, 221). Walters et al. (221) characterized *dcl* mutants unable to produce both clavulanic acid and cephamycin and demonstrated that complementation of the mutant strain with the *dclX* (*ccaR*) gene rescued production of both antibiotics. Perez-Llarena et al. (174) showed that mutation of the *ccaR* gene eliminated the biosynthesis of both cephamycin and clavulanic acid.

I have also investigated this regulatory gene in order to better understand its mechanism of regulation. In order to study the regulation of β-lactam biosynthesis by the CcaR protein it was necessary to create mutants of *S. clavuligerus* which lacked the CcaR protein. *ccaR* mutant strains that contained a disrupted or deleted *ccaR* gene but were otherwise isogenic to wildtype *S. clavuligerus* were generated by gene replacement via homologous recombination. All of the gene targeting constructs used for the gene replacement experiments were prepared as parallel sets in which the antibiotic resistance marker was inserted in both orientations relative to the target gene(s) being disrupted or deleted. This made it possible to determine if the mutation to the target gene(s) was responsible for the altered phenotype or if the antibiotic resistance marker had caused a polar mutation affecting other genes within the operon. This precaution regarding polar mutations in the *cob* genes (*ccaR*, *orf11* and *blp* genes) may be unnecessary since, the *ccaR* and *blp* genes were expressed as monocistronic transcripts and a transcript for *orf11* was undetectable (174).

Insertion of the *apr* cassette into the *ccaR* gene created the *ccaR*::*apr* disruption mutants. Deletion of DNA fragments containing the *ccaR* gene or the entire set of *cob* genes created the Δ*ccaR*::*tsr* and Δ*cob*::*tsr* mutants, respectively. All of the *ccaR* mutant strains generated by gene replacement lost the biosynthetic ability for cephamycin production in all media tested. These *ccaR* mutant strains have also lost the ability to produce clavulanic acid and clavams when grown in media specifically designed for their production, thus confirming an essential role for the CcaR protein in cephamycin and clavulanic acid production and also demonstrating that CcaR was required for clavam production. This makes the CcaR protein unique in that it simultaneously regulates the biosynthetic activity of three distinct biosynthetic pathways. Typically, multiple pathways within a single organism are each regulated by a pathway-specific transcriptional regulator, as in the case of *S. coelicolor* where RedD and ActII-ORF4 regulate RED and ACT biosynthesis, respectively (74, 155). *afsR* (77) and *bldA* (132) are global regulatory genes for antibiotic production within *S. coelicolor* that simultaneously control both undecylprodigiosin and actinorhodin biosynthesis but neither of these genes are located within the biosynthetic cluster.

The biosynthetic pathways responsible for DNR (DnrI, 206) and RED (RedD, 155) production in other *Streptomyces* spp. are known to be regulated by proteins homologous to CcaR. Mutants lacking the DnrI or RedD transcriptional activator displayed an antibiotic non-producing phenotype and lacked the transcripts responsible for the biosynthetic enzymes (140, 155).

Since homologous pathway-specific activators control antibiotic biosynthesis by regulating transcription of antibiotic biosynthetic genes, *ccaR* mutants were characterized for the synthesis of selected enzymes required for cephamycin biosynthesis. As expected, *ccaR* mutants were unable to make the CcaR protein and, as a result, they were also shown to be unable to make the LAT, δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase (ACVS), isopenicillin N synthase (IPNS) or desacetoxycephalosporin C synthase (DAOCS) proteins by western blot or SDS-PAGE analysis. Enzyme assays of cell-extracts prepared from *ccaR* mutants demonstrated that these strains completely lacked any ACVS, IPNS, isopenicillin N epimerase (IPNE) or DAOCS activity thus explaining the defect in cephamycin production.

Therefore the CcaR protein was essential to cephamycin production by activating production of at least these pathway enzymes. In most instances, the transcriptional activators like CcaR regulate a number of biosynthetic genes present within the cluster (23). Since LAT, ACVS and IPNS are believed to arise from a tricistronic operon, CcaR could potentially regulate cephamycin biosynthesis simply by controlling expression from the *lat* promoter. Similarly, IPNE and DAOCS, are enzymes from the middle of the cephamycin pathway, which arise from a separate polycistronic transcript which includes *cefD*, *cefE* (123) and possibly *pcd*, *cmcT* and *pbpA*. The fact that IPNE and DAOCS are also not produced in *ccaR* mutants indicates that the CcaR protein must regulate expression from a number of separate promoters. The late genes (*cefF*, *cmcH*, *cmcI* and *cmcJ*) are also believed to be transcribed as a polycistronic transcript and expression of this transcript may also be CcaR-dependent.

The *pcbC* gene can also be transcribed as a monocistronic transcript from a promoter located within the *pcbAB* coding region (178) and yet no IPNS protein was observed and no IPNS activity was detected in the *ccaR* mutants. This suggests that CcaR also regulates transcription of the *pcbC* promoter. However, promoter probe analysis of the *lat* and *pcbC* promoters using a high copy number promoter probe vector demonstrated that the *pcbC* promoter does not display the dramatic species-specific activation observed for the CcaR-dependent *lat* promoter (176). This suggested that expression of the *pcbC* promoter is not CcaR-dependent and therefore it seems unlikely that the lack of IPNS protein production is due to CcaR-dependent regulation of *pcbC* promoter expression. Rather it seems more likely that the lack of IPNS was related to the absence of IPNS protein observed in the Δ*lat*::*tsr*/term mutants. Disruption of the *lat*/*pcbAB*/*pcbC* polycistronic transcript should not affect transcription of the *pcbC* monocistronic transcript yet the Δ*lat*::*tsr*/term mutants were unable to produce any IPNS protein just like the *ccaR* mutants. The nature of the CcaR-dependent regulation of the IPNS protein production could be determined by northern analysis or S1 nuclease transcript analysis of the *pcbC* gene in RNA isolated from a *ccaR* mutant.

Presumably transcriptional regulation of the genes encoding biosynthetic enzymes also occurs in the clavulanic acid and clavam biosynthetic pathways to prevent their biosynthesis. Transcription of *claR*, the regulator responsible for expression of the late genes of clavulanic acid biosynthesis, was demonstrated to be absent in the *ccaR*::*apr* mutant (167). Mutants of *claR* accumulate clavaminic acid because the late genes, responsible for the conversion of clavaminic acid to clavulanic acid, are not expressed (165) but the *ccaR* mutants do not demonstrate an accumulation of clavaminic acid or any other imidazole-reacting intermediate in either SA or SOY media. Phenotypically, the *ccaR* mutants appear to be more like *orf2*::*apr* mutants, which also do not produce any imidazole-reacting intermediates including clavaminic acid, clavulanic acid or the clavams in either SA or SOY media (67). Mutants of the *orf2* gene suggest that its involved in one of the first steps of clavulanic acid biosynthesis. Duplicate copies or homologs exist for a number of the early clavulanic acid biosynthetic genes (ORFs 3-6) and these homologs are postulated to function in clavam biosynthesis (152, 168). Mutants of these early clavulanic acid biosynthetic genes are able to produce clavulanic acid in SOY medium because these clavam biosynthetic homologs compensated for the absence of the clavulanic acid homolog. The lack of clavulanic acid production in *orf2*::*apr* mutants grown in SOY medium suggests that the clavam biosynthetic cluster does not have a functional *orf2* homolog. The clavulanic acid and clavam biosynthetic pathways have been shown to share pathway intermediates up to the point of clavaminic acid (66). Thus a lack of *orf2* expression in a *ccaR* mutant would explain how both clavulanic acid and clavam production are blocked in a *ccaR* mutant. Therefore expression of the clavam biosynthetic genes may not be CcaR-dependent as the clavam negative phenotype would suggest but rather the lack of clavam biosynthesis could arise from the loss of *orf2* expression. Again, exactly how regulation of the clavulanic acid and clavams biosynthetic genes occurs in *ccaR* mutants would require transcript analysis of RNA isolated from the *ccaR* mutants.

Elimination of cephamycin biosynthesis in the *ccaR* mutants occurs due an absence of the production of the cephamycin biosynthetic enzymes. Without directly doing transcript analysis it was possible to demonstrate that the *ccaR* mutants lack transcription of the genes encoding cephamycin biosynthetic enzymes and conclude that CcaR-dependent regulation occurs at the level of transcription. Promoter probe analysis of wildtype *S. clavuligerus* and *ccaR* mutants transformed with a plasmid construct which contains the *lat* promoter and much of the *lat* gene in a single copy promoter probe vector showed a dramatic decrease, over 50-fold, in catechol 2,3 dioxygenase (C23O) specific activity in *ccaR* mutants as compared to the wildtype. In *ccaR* mutants the C23O activities fell below the basal level of activity provided by the promoter probe vector alone which suggests a very strong regulatory relationship between the transcriptional activator and the CcaR-dependent *lat* promoter. This decrease in promoter probe activity mirrors the species-specific expression that was observed by Petrich et al. (176) for the *lat*up promoter fragment in *S. lividans*. The lack of expression from the *lat* promoter in the *ccaR* mutants would eliminate transcription of the *lat/pcbAB/pcbC* polycistronic transcript which would directly result in the loss of LAT, ACVS and IPNS production.

In the absence of a gel-shift assay demonstrating the CcaR protein binding directly to DNA it was not possible to conclude that CcaR directly regulates transcription of these genes by binding to the promoters of CcaR-dependent genes. Attempts to use *S. clavuligerus* cell-extracts or the purified pMAL-CcaR fusion protein in gel-shift assays of the *lat* promoter were unsuccessful (data not shown). The CcaR-dependent phenotype may occur through a second regulatory protein or sigma factor whose production or at least activity is directly regulated by CcaR. However, the CcaR homolog DnrI has been shown to bind directly to DNA fragments containing DnrI-regulated promoters (209) and it is assumed that the homologous activators function in the same manner.

The CcaR-dependent production of the LAT protein in the Δ*ccaR*::*tsr* mutant was alleviated by replacing the *lat* promoter with the *ermE**-based constitutive high level expression cassette. Transformation of the *ermE*-lat* gene construct into the *ΔccaR::tsr* mutant via the single copy pSET152 integrating vector demonstrated that in the absence of the *lat* promoter production of the LAT protein was no longer CcaR-dependent. The *ccaR* mutant however remained unable to produce cephamycin because the IPNS protein and the other biosynthetic enzymes were still CcaR-dependent and therefore were not produced.

The addition of high copy number plasmids containing the *lat* promoter or the *lat* gene would greatly increase the number of activator recognition sites within the cell which should titrate the CcaR regulator. This would result in decreased activation of transcription from the chromosomal CcaR-dependent cephamycin biosynthetic promoters and therefore reduced cephamycin biosynthesis. For example, when a promoter fragment that was dependent on the WhiG sigma factor for expression was cloned into the high copy number promoter probe plasmid pIJ4083 and transformed into *S. coelicolor* sporulation was inhibited (208). The sporulation-deficient phenotype occurred as the multiple copies of the WhiG-dependent promoter caused a titration of the limiting amount of the WhiG protein.

Attempts to titrate the regulator of cephamycin biosynthesis were unsuccessful as the addition of the *lat* promoter or the *lat* gene were unable to significantly reduce cephamycin biosynthesis. How *S. clavuligerus* compensates for the increased number of recognition sites without neglecting the cephamycin biosynthetic cluster is unknown. The extra copies of the *lat* gene did however result in increased LAT protein production (data not shown) but whether the elevated *lat* expression was CcaR-dependent or driven by readthrough transcription from vector promoters was not known. The addition of a terminator in front of the cloned *lat* gene would eliminate transcription readthrough from vector promoters and demonstrate if the high level LAT production was CcaR-dependent. If the increased LAT production was CcaR directed then it would be important to observe the levels of the other biosynthetic enzymes and most importantly the CcaR protein and determine how these proteins respond to the addition of multiple activator recognition sites. Addition of the *dnrI* promoter region on a high copy number plasmid was also unsuccessful in causing a decrease in DNR biosynthesis by titrating out the DnrN regulator protein (83). The addition of multiple copies of the *redD* gene resulted in increased levels of the *redZ* transcript because of decreased negative autoregulation (255). The RedZ protein down-regulates expression at the *redZ* promoter and the increased number of RedZ-dependent promoter sites decreased the ability of the RedZ protein to regulate transcription at the *redZ* promoter.

CcaR-dependent expression was observed for the LAT, ACVS, IPNS, IPNE and DAOCS proteins. Absence of the early enzymes LAT, ACVS and IPNS arising from a tricistronic operon (176) and the middle enzymes IPNE and DAOCS which arise from a separate polycistronic transcript (123) indicates that CcaR must regulate expression from a number of separate promoters. Western blot analysis of a *S. clavuligerus* time course shows that the CcaR protein precedes production of the biosynthetic enzymes and that expression of the DAOCS protein was delayed relative to the LAT and IPNS proteins. That the CcaR protein was present before the other proteins conforms to the idea that the transcriptional activator CcaR was required for production of the biosynthetic enzymes. The question is how does the CcaR protein discriminate between early, middle and late genes with respect to the timing of expression? In the clavulanic acid pathway the late genes required for the modification of clavaminic acid are regulated by ClaR in a CcaR-dependent manner (165). The cephamycin gene cluster does appear to contain a second activator responsible for the expression of the middle and late genes and therefore the CcaR protein itself must be able to distinguish between early, middle and late promoters.

The ability to complement a *ccaR* mutant has been demonstrated previously as the *dclX*/*ccaR* gene cloned at a high copy number was able to restore antibiotic production (174, 221). The present study demonstrated that the Δ*ccaR*::*tsr* deletion mutant could be complemented, to return cephamycin biosynthesis, by a single copy of the *ccaR* gene *in trans* delivered by the integrating vector pSET152. The levels of cephamycin production seen in complemented mutants were reduced compared to wild type but this may be due to postional effects since pSET152 constructs integrate at the ØC31 *attB* site. Alternatively, some upstream regulating elements needed for optimum expression of *ccaR* may have been unknowingly omitted from the complementing fragment used. Extension of the sequence upstream of the *ccaR* gene has been reported to give higher levels of complementation (19). The successful complementation rescued cephamycin biosynthesis by returning production of LAT, IPNS, DAOCS and any other cephamycin biosynthetic enzyme that was expressed in a CcaR-dependent manner.

The placement of stop codons within the *ccaR* gene created mutant alleles that were unable to complement the *ccaR* mutant, showing that the *ccaR* gene does not contain alternate start codons capable of creating truncated but still functional versions of the CcaR protein. Since neither of these mutant *ccaR* constructs nor the C-terminal truncated constructs were able to complement the *ccaR* mutant, it was concluded that essentially complete "CcaR" gene products were needed before any complementation could occur. Modification of the final two CcaR amino acids in one of the complementation constructs created a CcaR protein that was barely detected by western blot analysis and was only capable of poorly activating cephamycin biosynthesis. Except for this minimally truncated gene, other truncated or prematurely terminated versions of the *ccaR* gene were unable to produce a stable truncated CcaR protein. Unfortunately most of the truncated versions also have additional non-*ccaR* encoded amino acids at the C-terminus before a stop codon was reached. Ideally, the truncated versions of CcaR would have been generated by the insertion of a stop codon directly into the *ccaR* gene by site-directed mutagenesis.

Since the anti-CcaR antibodies used in these studies are polyclonal in nature it seems unlikely that removing a portion of the C-terminus would create a protein that would not be detected by western analysis. The lack of truncated CcaR proteins may result from the instability of the truncated protein causing the protein to be degraded and removed from the cell. Alternatively, the CcaR protein may be positively autoregulated and the presence of a functional CcaR protein may be necessary for expression from the *ccaR* promoter. The mutant *ccaR* allele from NTG 83 containing the mutation converting proline 55 to a serine produced a non-functional CcaR protein that was unable to stimulate transcription of the cephamycin biosynthetic genes in the Δ*ccaR*::*tsr* mutant. The presence of this mutant protein suggests that either CcaR is not autoregulatory or that this particular protein is capable of autoactivating its own expression but can not activate expression of the cephamycin biosynthetic genes. Transcript analysis of *ccaR* expression in RNA isolated from the *ccaR*::*apr* mutants will determine whether CcaR autoactivates its own expression.

Four cephamycin non-producing NTG mutants were initially characterized and three of these mutants displayed a *ccaR* mutant phenotype. Of these mutants; NTG 26, NTG 32 and NTG 83 do not produce the CcaR protein or the LAT, ACVS, IPNS, IPNE and DAOCS biosynthetic activities and/or proteins or any of the β-lactam metabolites. The NTG 19 mutant was shown elsewhere to be cephamycin non-producing due to a mutation affecting the activity of the ACVS protein (166) and its lack of a *ccaR* mutant phenotype was confirmed. In addition, a *dclX*68 mutant generated by random mutagenesis, and defective in CcaR/DclX production, was also obtained (221). The three NTG mutants and the *dclX*68 mutant were further characterized. Southern analysis suggested that none of these mutants contained a large deletion within the *ccaR* gene and they were presumed to be point mutants. The *ccaR* alleles were cloned from each mutant and sequenced. The *ccaR* alleles of *dclX*68 and NTG 32 were shown to have mutations that generate stop codons resulting in premature termination of the CcaR protein. The premature termination of the CcaR protein prevented either of these *ccaR* alleles from returning production of cephamycin or the CcaR protein to the Δ*ccaR*::*tsr* mutant. The

NTG 83 *ccaR* allele contained a mutation that caused the conversion of proline 55 to a serine residue. This mutation created a gene that was unable to complement the Δ*ccaR*::*tsr* mutant to produce cephamycin or any of the biosynthetic enzymes but was capable of producing a non-functional CcaR protein. Often proline residues effect the protein chain conformation to play a role in protein folding and therefore the final protein structure (138). The effect of the two adjacent proline residues makes it more likely that the proline to serine change alters the folding of the NTG 83 CcaR protein and prevents it from functioning as a transcriptional activator. Protein secondary structure predictions done using the UWGCG package suggested that the mutant CcaR protein may adopt a different conformation than the wildtype CcaR protein (data not shown). When the Garnier, Osguthorpe and Robson algorithm (85) was used the mutant and the wildtype CcaR proteins had two different predicted secondary structures, while the Chou-Fasman algorithm (45, 46) within the same program suggested that the mutation would not alter the conformation of the mutant CcaR protein.

The NTG 26 mutant *ccaR* allele did not have any mutations in the coding region or the promoter sequence upstream of the *ccaR* gene. Addition of this *ccaR* allele to the Δ*ccaR*::*tsr* mutant rescued cephamycin biosynthesis but at a reduced level compared to the wildtype *ccaR* gene. The NTG 26 allele was inserted into pSET152 in the opposite orientation to the wildtype *ccaR* allele and this may have affected its level of expression.

Western analysis was unable to detect a CcaR protein in NTG 26 or NTG 83 yet both of these mutants contain alleles that have demonstrated the ability to return production of the CcaR protein to the Δ*ccaR*::*tsr* mutant. The Δ*ccaR*::*tsr* mutant created by homologous recombination was otherwise isogenic to wildtype *S. clavuligerus* and this suggests the presence of a secondary mutation(s) within both NTG 26 and NTG 83 which is preventing expression of the *ccaR* gene.

Transformation of the NTG 26 mutant with a second copy of the wildtype *ccaR* gene via an integrating vector resulted in low level expression of the biosynthetic enzymes as judged by western analysis but no cephamycin production. Transformation with the a *ccaR* construct in which the normal *ccaR* promoter was preceded by an *ermE** promoter greatly increased production of the cephamycin biosynthetic enzymes and resulted in low level production of cephamycin. The low levels of cephamycin production observed for these tandem promoter *ccaR* construct transformants suggests that NTG 26 may possess another mutation within one of the biosynthetic enzymes that decreases but does not eliminate activity, resulting in decreased cephamycin levels. Analysis of the ability of the tandem promoter *ccaR* construct to return clavulanic acid production to the NTG 26 transformants may confirm that a mutation has occurred to a cephamycin biosynthetic gene. The low levels of complementation observed for the biosynthetic enzymes in wildtype *ccaR* construct transformants which already contain a wildtype copy of the *ccaR* gene in their genomes, may have resulted from transcriptional readthrough of the introduced second copy of the *ccaR* gene from vector or chromosomal promoter(s) rather than authentic expression from the *ccaR* promoter. This transcriptional readthrough may have generated a small number of CcaR molecules which resulted in low level stimulation of the cephamycin biosynthetic enzymes. The addition of a terminator in front of the *ccaR* gene would allow this observation to be tested. Unfortunately the NTG 83 mutant, which also appears to lack expression of the *ccaR* allele could not be transformed with the integrating vectors. As an alternative approach this lack of CcaR production in the mutants, which contain *ccaR* alleles capable of producing the CcaR protein in the Δ*ccaR*::*tsr* background, could be examined by transcript analysis of RNA isolated from wildtype *S. clavuligerus* as well as the NTG 26, NTG 32 and NTG 83 mutants. If the NTG 26 and NTG 83 mutants are negative for *ccaR* expression it would suggest that along with their other mutations these strains also lack a transcriptional activator responsible for the activation of CcaR. Transcriptional regulators for both DnrI (DnrN, 83) and RedD (RedZ, 225) exist within the DNR and the RED biosynthetic clusters, respectively. The disruption of either *dnrN* or *redZ* completely eliminates transcription of *dnrI* or *redD*, respectively and as a result eliminates transcription of the biosynthetic enzymes. Either of the NTG mutants could be a mutant for a DnrN/RedZ homolog but no such candidate gene was located within the cephamycin biosynthetic cluster.

The NTG 32 mutant behaves like the Δ*ccaR*::*tsr* mutant as complementation with the wildtype *ccaR* gene or the tandem promoter *ccaR* construct result in strong production of cephamycin.

Attempts to replace the *ccaR* promoter with the strong constitutive *ermE**-based expression cassette at an artificial *Nde*I site created at the ATG start codon predicted by Walters et al. (221) were unsuccessful. The transformants generated by introduction of the *ermE*-ccaR* fusion construct grew slowly on solid media and very poorly in liquid media. The toxicity of this construct was not attributed to cephamycin overproduction since the transformation of the Δ*lat*::*tsr* mutant, which can not produce cephamycin, with this construct was equally unsuccessful. It is possible that the CcaR GTG start codon predicted by Perez-Llarena et al. (174) is the authentic start codon for CcaR and the production of a CcaR protein with six extra amino acids was toxic to the cell. Creation of a *Nde*I site at the other possible *ccaR* gene start codon would allow this hypothesis to be tested. Overexpression of the *dnrI* gene by replacing the native promoter with *ermE** promoter also resulted in inhibition of growth of the wildtype *S. peucetius* strain (206).

Use of the *dpsB* mutant strain which would not produce the toxic intermediates of daunorubicin biosynthesis allowed overproduction of an active DnrI protein (209). Therefore in *S. clavuligerus* it seemed likely that the overproduction of a CcaR-regulated metabolite other than cephamycin was responsible for the toxicity rather than the toxic accumulation of a mutant CcaR protein.

The presence of the *orf11* and *blp* genes in the cephamycin cluster was not predicted and an obvious role for these genes in antibiotic biosynthesis was unknown. Since all of the enzyme activities required for cephamycin biosynthesis had been accounted for it was presumed that these genes were more likely to play a role in regulation or resistance. Database searches provided little insight towards the function of the Orf11 protein as the protein displayed no homology to any known proteins. Heterologous probing detected homologs of the *orf11* gene in *S. jumonjinensis* and *S. lipmanii* but not in any other β-lactam producers. These strains and *S. clavuligerus* are similar in their ability to produce both cephamycin and clavulanic acid or clavam-like metabolites (69). *S. lipmanii* likely does not produce clavulanic acid or the clavams but production of precursors such as clavaminic acid and proclavaminic acid have been detected along with clavaminate synthase activity (19, 69). Heterologous probing also detected homologs of the *cla* gene (now referred to as *pah*, 95) coding for the early gene proclavaminic acid amidino hydrolase in both *S. jumonjinensis* and *S. lipmanii* (6), thus confirming the relationship between these strains and the biosynthesis of clavam-like metabolites.

The *orf11* gene along with the rest of the *cob* genes was deleted and replaced with the *tsr* marker to create the Δ*cob*::*tsr* mutants. The loss of antibiotic production in the Δ*cob*::*tsr* mutants was predictable due to the lack of a functional *ccaR* gene and any effects that would be caused by the absence of *orf11* were obscured by the overriding effect of the *ccaR* deletion. Complementation of the Δ*cob*::*tsr* mutant with the wildtype *cob* fragment rescued production of all of the β-lactam metabolites. Complementation of the Δ*cob*::*tsr* mutant with a construct containing the *cob* DNA fragment with the *orf11* gene mutated by insertion of the translational stop fragment (TSF) also returned production of all of the β-lactam metabolites. Only the *cob* gene construct containing the mutated *ccaR* was unable to complement the Δ*cob*::*tsr* mutant to return production of the metabolites. Due to the small size of the translational stop fragment (37 bp) its presence within the target gene should not negatively effect transcription of downstream genes within an operon. This precaution may have been unnecessary since a previous study demonstrated that *ccaR* and *blp* were transcribed as monocistronic transcripts and no transcript was detected for *orf11* in trypticase soy broth (174). The lack of a *orf11* transcript can be explained if *orf11* is indeed involved in clavulanic acid or clavam biosynthesis. In that case, expression of the *orf11* gene, like the clavulanic acid and clavam biosynthetic genes, may only occur in specific media and in *S. clavuligerus* clavulanic acid production is known to be poor in trypticase soy broth.

Disruption mutants of the *orf11* gene were also constructed via homologous recombination by replacement with a copy of *orf11* disrupted by insertion of the *apr* cassette. Disruption of the *orf11* gene did not eliminate production of any of the metabolites regardless of the orientation of the *apr* cassette. This eliminates an essential role for *orf11* in the production of any of the metabolites. The decrease in cephamycin production observed in the *orf11*::*apr*B mutants (*apr* oriented in the opposite direction compared to *orf11*) did not suggest that *orf11* was a regulator of cephamycin biosynthesis since a similar decrease was not observed in the *orf11*::*apr*A mutants or the Δ*cob*::*tsr* mutant complemented with an *orf11* mutated *cob* fragment.

Despite this inability to demonstrate an essential role for *orf11* in the production of any of the three classes of metabolites examined, it is still tempting to suggest that it may play a role in either clavulanic acid or clavam biosynthesis because of the cross-hybridization in *S. jumonjinensis* and *S. lipmanii*. Perhaps *orf11* is not essential for clavulanic acid or clavam biosynthesis but mutation of the gene could increase or decrease the amount of metabolite made or possibly *orf11* functions under conditions which are difficult to duplicate in the laboratory.

The *blp* gene is located adjacent to *orf11* and comprises the third gene of the *cob* group of genes. Database homology searches list the β-lactamase inhibitory protein (BLIP), a proteinaceous β-lactamase inhibitor from *S. clavuligerus* (62) as the only homolog to the Blp protein. Any functions for the BLIP protein in antibiotic biosynthesis other than its obvious β-lactamase inhibition are unknown. Whether the Blp protein can function as a β-lactamase inhibitor is unknown but seems unlikely as many of the critical residues for BLIP and TEM β-lactamase interaction are not conserved in Blp. The *blp* gene does not contain a homologous gene in any of the other β-lactam producing or non-producing strains. Genes homologous to the *bli* gene were also not detected by heterologous probing (62).

It was not possible to create a disruption mutant of the *blp* gene by homologous recombination in a parental wildtype background capable of producing antibiotic. The inability to produce the *blp*::*apr* mutants in the wildtype background was similar to the attempts to mutate the *pcbR* gene which is apparently necessary for survival in the presence of the high levels of cephamycin produced by the wildtype strain (166). However, the *blp* gene was deleted along with the rest of the *cob* genes and replaced with the *tsr* marker to create the Δ*cob*::*tsr* mutants. The loss of antibiotic production in the Δ*cob*::*tsr* mutants was due to the lack of a functional *ccaR* gene. With the Δ*cob*::*tsr* mutant it was possible to determine if *blp* played a role in antibiotic production by complementing the Δ*cob*::*tsr* mutant with a construct which contains the *cob* DNA fragment with the *blp* gene mutated by the insertion of the stop codons. The ability of the Δ*cob*::*tsr* mutant to be complemented with that construct suggested that Blp was not involved in resistance and that it should have been possible to generate *blp*::*apr* mutants in a wildtype background. Complementation of the Δ*cob*::*tsr* mutants with the *blp*-mutated *cob* fragment was also successful in returning production of the metabolites. This eliminates an essential role for *blp* in the production of any of the metabolites.

VI.3 Analysis of *lat*, *pcbAB* and *pcbC* regulation

The *lat* gene, present exclusively in prokaryotic β-lactam producers, represents the first step of cephamycin biosynthesis (142). Since the *lat*, *pcbAB* and *pcbC* genes are separated by small intergenic regions and have the same transcriptional orientation the genes are likely organized in an operon that coordinates expression of the three genes (170). Petrich et al. (176) suggested that the *lat* promoter was responsible for initiating a tricistronic transcript of approximately 14 kb which contains the *lat*, *pcbAB* and *pcbC* coding regions. The co-transcription of these steps provided a logical means to coordinate control over the onset of antibiotic production (175).

The "co-transcription" model proposed for the regulation of expression of these genes may be an oversimplification as other factors which may effect transcription of these genes have been discovered. A potential stem-loop structure is located within the 152 bp intergenic region between the *lat* and *pcbAB* coding regions (215, 234) and this structure was preceded by a putative anti-terminator region proposed to regulate the stem-loop structure's ability to function as a transcription terminator (234). Yu et al. (234) also demonstrated the presence of a promoter region within the *lat* coding region and suggested that it was capable of driving *pcbAB* transcription. As well, the *pcbC* promoter located within the 3' end of the *pcbAB* coding region created a 1.2 kb *pcbC* monocistronic transcript (178).

Piret and coworkers (56, 181, 234) suggested that the *lat*, *pcbAB* and the *pcbC* genes had an "interdependence" upon each other for their transcription. αAA, the product of LAT and PCD activities, would stimulate *pcbAB* transcription to produce ACVS. The ACVS protein would catalyze production of the δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV) tripeptide and that would in turn stimulate *pcbC* transcription to produce IPNS. This suggested that the genes would give rise to three transcripts originating from the *lat*, putative *pcbAB* and *pcbC* promoters and that the transcription of each would be dependent upon the action of the preceding gene product for their transcription.

Before promoter probe analysis was initiated in this study, the promoter probe vectors were redesigned in an effort to address some shortcomings of the high copy number promoter probe plasmid pIJ4083. Since pIJ4083 contains only a *Streptomyces* replicon, construction of recombinant promoter probe vectors was very time consuming, compared to similar manipulations carried out in an *E. coli* background. Therefore, the *E. coli/Streptomyces xylE*-based promoter probe shuttle vectors pDA2000 and pDA2500 were constructed. This allowed the recombinant promoter probe constructs to be generated quickly in *E. coli* before the final constructs were moved into *S. clavuligerus* for analysis.

There were also concerns about the moderate basal levels of C23O activity which can arise from pIJ4083, fragments that do not necessarily have a functional promoter activity under physiological conditions nonetheless give strong activity as in the case of the putative *pcbAB* promoter (169). Finally, certain regulatory factors may be titrated out by the high copy number plasmid altering promoter strength and regulation from what would occur normally *in vivo* (175). Therefore efforts were made to greatly decrease the copy number of the promoter probe vectors with hopes of decreasing these concerns.

The mid-copy number promoter probe plasmid pDA2500 uses the pJV1-based origin of replication in *Streptomyces* while the single copy promoter probe vector pDA2000 uses the *att* fragment containing the ØC31 attachment site (*attP*) and integrase from pSET152 (27) to integrate into the chromosome of *S. clavuligerus* (126). The pDA2000 and pDA2500 vectors themselves did not display any instability problems in *Streptomyces* but problems did occur during vector construction when strong promoters that were active in *E. coli* were inserted.

The presence and/or function of the *lat* transcription terminator and the putative *pcbAB* promoter were tested by promoter probe plasmid analysis using the single copy promoter probe vector pDA2000. Function of the transcription terminator was determined by comparing the C23O specific activities in wildtype *S. clavuligerus* transformed with two different constructs. One plasmid construct contained the *lat* gene but would not have a functional *lat* stem-loop structure and the second plasmid construct contained both the *lat* gene and the intact downstream stem-loop. No dramatic differences in C23O specific activities were detected suggesting that the stem-loop structure was not functioning as a terminator for *lat* transcription.

The lack of termination at the *lat* stem-loop was also demonstrated by the inability to detect a transcription stop point within the *lat-pcbAB* intergenic region by S1 nuclease transcript analysis (175). As well, northern blot analysis did not detect a monocistronic *lat* transcript but rather a smear of degraded high molecular weight mRNA which was very similar to the results obtained from attempts to detect a *pcbAB*-containing transcript (176). Yu et al. (234) suggested that the LAT protein may serve another unrecognized function in *S. clavuligerus* and that regulation at the *lat* stem-loop structure would allow monocistronic *lat* expression while preventing the expression of *pcbAB* and *pcbC*. Termination at the *lat* stem-loop would not likely occur during the optimal antibiotic production conditions that result from growth in trypticase soy broth plus 1% (w/v) starch (TSBS). The promoter probe analysis done in this study and the mRNA analysis done by Petrich et al. (176) were both done during conditions optimal for antibiotic production. Therefore conditional regulation of termination at the *lat* stem-loop structure can not be ruled out but these experiments confirm that the *lat* stem-loop does not impede transcription of the polycistronic transcript under optimal antibiotic production conditions.

The transcription terminator downstream of *pcbC* resulted in an 8-fold decrease in apparent promoter activity as compared to the *pcbC*up promoter fragment in the promoter probe plasmid pIJ4083 (178). The *lat* stem-loop structure was calculated to have a strength of -53.8 Kcal per mol using the method of Salser (193) while the *pcbC* gene was predicted to contain as many as four stem-loop structures (63, 109) Some of these stem-loop structure may terminate the *pcbR* gene which also terminates in this intergenic region. The *pcbC* stem-loop structures were predicted to have a calculated strengths of -39.3, -49.6, -12.6 and -56.1 Kcal/mol, again using the method of Salser (193). The fact that transcription proceeds through the *lat* stem-loop structure opens the possibility that the stem-loop does have an antiterminator regulation relationship (234) and this allows transcription to proceed through the stem-loop. Alternatively, the stem-loop may not be a terminator but the structure may protect the *lat* coding region part of the transcript from degradation.

Yu et al. (234) also characterized a promoter region present within the *lat* coding region that was proposed to initiate *pcbAB* transcription. A DNA fragment containing the putative promoter(s) displayed very strong promoter activity in pIJ487, a promoter probe plasmid which uses the promoterless neomycin resistance marker as the reporter of promoter activity. This region also demonstrated strong promoter activity in the *xylE*-based promoter probe plasmid pIJ4083 (169).

The function of the putative *pcbAB* promoter was also judged by promoter probe analysis using the single copy promoter probe vector pDA2000. A series of promoter probe constructs were prepared each containing progressively more DNA from the *lat* coding region. Two constructs contain 0.2 and 1.1 kb of DNA from the *lat* gene but neither contain the putative *pcbAB* promoter region while the two remaining constructs contain 1.6 kb and 1.8 kb of DNA respectively, which includes the *pcbAB* promoter region. Comparison of these four constructs demonstrated that the inclusion of progressively larger fragments of the *lat* gene including those containing the putative *pcbAB* promoter region did not enhance C23O specific activity markedly suggesting that the *pcbAB* promoter activity was not detected during promoter probe analysis at single copy. The lack of promoter activity from the putative *pcbAB* promoter was also suggested by the inability to detect a transcription start point (*tsp*) within this region by both high resolution and low resolution S1 nuclease transcript analysis of RNA isolated from wildtype *S. clavuligerus* (111).

Promoter probe analysis suggested that the *lat* stem-loop structure and the putative *pcbAB* promoter appear to be physiologically non-functional under the conditions tested. The inability to demonstrate any promoter activity from the *pcbAB* promoter region suggested that under antibiotic producing conditions the transcription of the three genes was interdependent upon each other because they were part of a polycistronic transcript. The results generated from promoter probe analysis do not allow strong conclusions to be drawn in the absence of direct mRNA analysis. Therefore it was felt that transcriptional organization of these genes and the function of these other factors could best be studied *in vivo* by manipulation of the operon within the *S. clavuligerus* chromosome.

Jensen and coworkers (111) initiated this project with the creation of the *lat*::*apr* mutant strain. The apramycin resistance (*apr*) marker was inserted in the opposite orientation to the direction of transcription for the *lat* gene so that transcription from the *lat* promoter into the *pcbAB* and *pcbC* genes would be interrupted. The goal for the *lat*::*apr* mutant was to determine whether interruption of the polycistronic transcript originating from the *lat* promoter would eliminate ACVS activity if the putative *pcbAB* promoter was not physiologically functional. If the "interdependence" model was correct the disruption of the *lat* gene would eliminate ACVS production since the absence of αAA would prevent the *pcbAB* promoter from initiating transcription. The *lat*::*apr* mutant did produce ACVS activity but at reduced levels compared to wildtype (111). This could represent evidence that the putative *pcbAB* promoter was functional. However, S1 nuclease transcript analysis of RNA isolated from the *lat*::*apr* mutant was unable to determine where the *pcbAB* transcription was originating. It was also possible that a divergent promoter was located within the *apr* marker and was initiating *pcbAB* transcription rather than the putative *pcbAB* promoter or that transcripts originating from the *lat* promoter were managing to readthrough the divergently oriented *apr* marker.

Therefore three different sets of *lat* mutants were generated in this study and these *lat* mutants were designed to unequivocally determine the transcriptional organization of the *lat, pcbAB* and *pcbC* genes and conclude whether the putative *pcbAB* promoter was responsible for *pcbAB* transcription *in vivo*. The *lat* mutant strains were generated by gene replacement by inserting a modified *lat* gene into the *S. clavuligerus* chromosome. Since the *lat* mutant strains were created by homologous recombination they should be otherwise isogenic to wildtype *S. clavuligerus*. All of the gene targeting constructs used for generating partial *lat* deletions in wildtype *S. clavuligerus* were prepared as parallel sets in which the thiostrepton resistance (*tsr*) marker was inserted in both orientations relative to the target gene. However, gene replacement mutants were obtained for only one orientation of each of the Δ*lat*::*tsr*/term mutants and the Δ*lat*::*tsr* mutants. Fortunately, in both instances the *lat* mutants obtained had the *tsr* marker inserted in the opposite orientation relative to the biosynthetic genes and because the goal was to generate a polar mutation within the polycistronic message these were the desired mutants.

In the Δ*lat*::*tsr* partial deletion mutants, a DNA fragment of about 0.5 kb from the 5' end of the *lat* coding region was deleted and replaced with the *tsr* marker. The Δ*lat*::*tsr* mutants would be similar to the *lat*::*apr* mutant strain except that the resistance marker had been changed to *tsr* which allowed the mutant to be transformed with pSET152-based plasmids and part of the *lat* coding region had been deleted. In the second set of mutants, a larger fragment which included the *lat* promoter and 0.5 kb from the 5' end of the *lat* coding region was deleted and replaced with the *tsr* marker followed by the 31-demethyl-FK506-*O*-methyltransferase (FKMT) terminator which created the Δ*lat*::*tsr*/term partial deletion mutants. The ability of the FKMT terminator to stop transcription from the upstream *tsr* marker was demonstrated by promoter probe analysis. Therefore it was felt that the FKMT terminator would successfully eliminate readthrough transcription in the Δ*lat*::*tsr*/term mutants to prevent transcription from proceeding into the *pcbAB* and *pcbC* genes. To ensure the success of the FKMT terminator the *lat* promoter was also removed in these mutants. In both of these *lat* partial deletion mutant sets the putative *pcbAB* promoter was still present but transcriptional readthrough from upstream promoter sequences would be interrupted by the *tsr* marker or the transcriptional terminator.

The Δ*lat*::*tsr* mutants were unable to produce cephamycin even after 72 hours of growth in TSBS, however the addition of 2 mM αAA to the TSBS allowed very low levels of cephamycin to be produced. The Δ*lat*::*tsr*/term mutants were also unable to produce cephamycin in TSBS but unlike the Δ*lat*::*tsr* mutants the addition of αAA to the TSBS did not allow any cephamycin to be produced. These mutants were obviously unable to produce cephamycin due to the elimination of the essential LAT protein. The addition of αAA, which is the product of the LAT and PCD catalyzed reactions, compensated for the absence of the LAT protein and allowed the Δ*lat*::*tsr* mutants, which presumably produce ACVS and IPNS protein because of readthrough transcription, to produce cephamycin. The Δ*lat*::*tsr*/term mutants were also unable to produce the essential LAT protein but the addition of exogenous αAA did not compensate for the loss of the LAT protein.

Western blot analysis of the Δ*lat*::*tsr* and Δ*lat*::*tsr*/term mutants confirmed that these mutants were unable to produce the LAT protein. The mutants were also shown not to produce IPNS or ACVS protein but they were capable of producing near wildtype levels of DAOCS. The addition of αAA to the mutants did not appear to result in increased levels of IPNS protein. However, the Δ*lat*::*tsr* mutants must be able to produce undetectable amounts of ACVS and IPNS protein since the addition of αAA allowed the production of low levels of cephamycin. This conclusion was confirmed and quantified by showing that the Δ*lat*::*tsr*A1 mutant was able to produce IPNS activity but at only 14% of wildtype enzyme activity levels. The ACVS activity remained below detectable levels with the enzyme assay used but obviously it must also be present at low levels. The low level of IPNS activity detected must be below the levels of detection of the IPNS protein by western analysis. The Δ*lat*::*tsr*/termA1 mutant, which could not produce cephamycin when supplemented with αAA, did not have detectable ACVS or IPNS enzyme activities. This confirmed that the transcription terminator blocked the low levels of readthrough transcription observed in the Δ*lat*::*tsr* mutant. The Δ*lat*::*tsr*A1 and Δ*lat*::*tsr*/termA1 mutants however produced near wildtype amounts of IPNE and DAOCS enzyme activities.

It was demonstrated that increasing the amounts of exogenous αAA continued to increase the amount of cephamycin produced by the *lat*::*apr* mutant (111). Therefore the levels of intracellular αAA achieved via exogenous feeding are limiting cephamycin production by the *lat*::*apr* mutant and these amounts may also fall below the levels necessary for the stimulation of *pcbAB* expression. In an effort to increase the intracellular levels of αAA, each of the *lat* mutant sets was transformed with pSET152-based integrating vectors containing either the wildtype *lat* gene or the *lat* gene under the control of a strong *ermE**-based expression cassette.

Complementation with the wildtype *lat* gene or *ermE**-*lat* expression constructs allowed greater amounts of cephamycin to be produced but the values for the Δ*lat*::*tsr* mutants were still considerably lower than wildtype *S. clavuligerus*. Transformation with the negative control plasmid pDA1000 containing the *tsr* marker resulted in no cephamycin production. Complementation of the Δ*lat*::*tsr*/term mutants with the wildtype *lat* gene, *ermE*-lat* construct or the negative control construct, like the exogenous αAA feeding, was unable to return cephamycin production to these mutants. Western blot analysis confirmed that transformation with either of the *lat* gene constructs returned production of high levels of the LAT protein to both of the *lat* mutants. This expression or overexpression of the *lat* gene should result in wildtype or higher levels of αAA within the mutants and did increase cephamycin production yet the production of the ACVS and IPNS proteins did not increase. The presence or absence of a functional *lat* gene does not appear to effect DAOCS protein production.

Deletion of the *lat* promoter and the insertion of the FKMT terminator into the *lat* coding region would prevent transcription from proceeding into the *pcbAB* and *pcbC* genes. The Δ*lat*::*tsr*/term mutants demonstrated that termination of transcription within the *lat* gene prevented the accumulation of any IPNS or ACVS protein. Even the simple replacement of part of the *lat* coding region with the *tsr* marker in the Δ*lat*::*tsr* mutants was successful in almost completely eliminating production of the ACVS and IPNS proteins. Therefore the termination or the interruption of the *lat*/*pcbAB*/*pcbC* polycistronic transcript originating from the *lat* promoter directly affects the production of the ACVS and IPNS proteins. This also suggests that the *lat* stem-loop structure located between the *lat* and *pcbAB* coding regions does not prevent transcription from proceeding into the *pcbAB* and *pcbC* genes.

The amount of ACVS protein made by these *lat* mutants sets was not enhanced by the addition of the *lat* gene *in trans* or exogenous αAA. The complemented Δ*lat*::*tsr* mutants were only able to produce minimal amounts of cephamycin because of the low levels of ACVS and IPNS proteins and the complemented Δ*lat*::*tsr*/term mutants were unable to make any cephamycin because these mutants do not make any ACVS or IPNS protein. Since the putative *pcbAB* promoter region described by Yu et al. (234) was still present within the Δ*lat*::*tsr*/term mutants but did not lead to any ACVS production it appears that under these conditions the *pcbAB* promoter, in the presence or absence of αAA, was unable to direct transcription of *pcbAB*.

The Δ*lat* mutants were created by a second gene replacement procedure applied to the original *lat*::*apr* mutant strain such that the mutated *lat* gene containing the *apr* marker was replaced with a mutated *lat* gene that contains a deletion removing an internal fragment. The Δ*lat* mutants were designed to create a *lat* mutant which did not make a functional LAT protein but in which transcription of the polycistronic mRNA was not altered. It was hoped that the removal of this internal fragment would not greatly affect the stability of the transcript since attempts to repair the *lat*::*apr* mutant with a *lat* gene that contained only a frameshift mutation were unsuccessful. The loss of a functional LAT protein would demonstrate if ACVS and IPNS production are dependent upon the production of αAA. In the experiments designed to "repair" the *lat::apr* mutant strain the putative Δ*lat* mutants were screened for the loss of the apramycin resistant phenotype (apr^R) which is a considerably less powerful screening technique than selection for a new antibiotic resistance phenotype and therefore mutants were more difficult to obtain and/or detect. Again the region containing the putative *pcbAB* promoter was present in the Δ*lat* mutants.

The Δ*lat* mutants were also unable to produce cephamycin after 72 hours of growth in TSBS but the addition of αAA to the TSBS allowed the mutant strains to produce greater amounts of cephamycin than was produced by the parental *lat::apr* strain. The Δ*lat* mutants were unable to make cephamycin due to the elimination of the essential LAT protein but the addition of αAA compensated for the absence of the LAT protein.

Western blot analysis of the Δ*lat* mutants confirmed that the LAT protein was not produced but after longer exposures a faint band of approximately 35 KD was observed. The internal deletion which created the Δ*lat* mutants was in frame and should produce a protein of 37.2 KD and this faint band is possibly the deleted LAT protein which has decreased stability. The mutants were also shown to produce high levels of IPNS, DAOCS and ACVS protein. The addition of αAA to the Δ*lat* mutants or the *lat::apr* mutant did not appear to increase the levels of IPNS protein. The Δ*lat* mutants were quite capable of producing high levels of ACVS and IPNS which were much increased over the ACVS and IPNS levels in the parental *lat::apr* mutant. Although the deleted form of the LAT protein was inactive the Δ*lat*A1 mutant was able to produce the high levels of ACVS and IPNS activity in the absence of αAA. It appears that the internal deletion within the *lat* gene did not significantly affect the stability of the *lat/pcbAB/pcbC* polycistronic transcript and that production of αAA was not needed for the expression of *pcbAB* and *pcbC* genes.

Complementation of the Δ*lat* mutants with the wildtype *lat* gene or *ermE*-lat* constructs returned strong levels of cephamycin production to these mutants. Complementation of the Δ*lat* mutants with these *lat* gene constructs mirrored the αAA feeding results except that complementation allowed cephamycin to be produced to higher levels, confirming that expression of the *lat* gene created higher intracellular levels of αAA than could be achieved with the exogenous αAA feeding experiments. Complementation with the negative control plasmid pDA1000 resulted in no cephamycin production.

Western blot analysis confirmed that complementation with the *lat* gene constructs returned the production of high levels of the LAT protein to the Δ*lat* mutants. The high levels of ACVS and IPNS proteins produced by the Δ*lat* mutants allowed them to make high levels of cephamycin when the *lat* gene was added *in trans*. This suggested that the Δ*lat*::*tsr* mutants were limited in the amount of cephamycin produced when αAA was fed because of the low levels of ACVS and IPNS protein produced.

The three different sets of *lat* mutants were designed to determine whether the "interdependence" or the "co-transcription" model was regulating transcription of the *lat*, *pcbAB* and *pcbC* genes *in vivo*. The promoter probe analysis suggested that under these particular growth conditions neither the *pcbAB* promoter nor the *lat* stem-loop structure were functional and that the *lat* and *pcbAB* coding regions were present on the same transcript. The presence of these genes on a polycistronic transcript was supported by the Δ*lat*::*tsr*/term mutants which had the FKMT terminator inserted into the *lat* coding region well before the putative promoter region predicted to be responsible for *pcbAB* transcription. That these mutants did not produce any ACVS or IPNS protein could be explained by the fact that the mutants did not produce αAA which might be necessary for the stimulation of *pcbAB* transcription or it could be due to the fact that the genes are co-transcribed and the terminator disrupts the polycistronic transcript. When a functional *lat* gene was added to the Δ*lat*::*tsr*/term mutant *in trans* the mutant was still unable to make any cephamycin and ACVS or IPNS protein. That the addition of αAA to Δ*lat*::*tsr*/term mutant did not stimulate ACVS and IPNS production as would be predicted by the "interdependence" model demonstrated that the premature termination of a *lat*/*pcbAB*/*pcbC* polycistronic transcript eliminated production of the gene products located downstream of the terminator. Even the Δ*lat*::*tsr* mutants, which have the *tsr* marker inserted into the *lat* coding region, were almost unable to produce ACVS and IPNS activity and production of these proteins was not enhanced when a functional *lat* gene was added *in trans*. The *lat*::*apr* mutant also had reduced ACVS and IPNS activities but their levels were decreased to a lesser degree because either the *apr* marker was less effective in disrupting transcription or the *apr* marker fragment contained a divergent promoter.

Due to the high cost of synthetic ACV tripeptide it was not possible to feed the *lat* mutants with ACV to see if this would stimulate IPNS production as predicted by the interdependence model. It seems unlikely that ACV would stimulate IPNS production in the Δ*lat*::*tsr* or Δ*lat*::*tsr*/term mutants since the Δ*lat* mutants are quite capable of producing high levels of IPNS protein in the absence of αAA and therefore by definition, in the absence of ACV. When a functional *lat* gene was added to the Δ*lat* mutants the amounts of IPNS protein did not increase. However in preliminary studies underway at present, putative *pcbC* disruption mutants have been generated and these *pcbC* mutants should accumulate ACV since biosynthesis would be blocked at this step. Therefore the Δ*lat*::*tsr* or Δ*lat*::*tsr*/term mutants could be grown in fresh TSBS media supplemented with TSBS media containing ACV harvested from a *pcbC* mutant fermentation to determine if this would stimulate IPNS production. This requires that the accumulation of ACV in the cell will also lead to accumulation of ACV in the media.

Studies using the Δ*lat* mutants approached the *lat*, *pcbAB* and *pcbC* regulation question from a different direction and demonstrated that the ACVS and IPNS proteins could be produced in the absence of αAA. These Δ*lat* mutants demonstrated that maintenance of the polycistronic transcript even in the absence of a functional LAT protein allowed the ACVS and IPNS proteins to be produced at high levels. In fact the addition of the *lat* gene *in trans* did not elevate the levels of either the ACVS or IPNS proteins.

The fact that the *lat* promoter was equally active in the presence or absence of αAA was confirmed by promoter probe analysis. The *lat* promoter fragment in a single copy promoter probe construct was used to compare the strength of the *lat* promoter in wildtype *S. clavuligerus* versus its strength in the various *lat* mutants to determine if αAA plays a role in stimulating transcription from the *lat* promoter. Promoter probe analysis suggested that the *lat* promoter was equally active in the presence or absence of αAA as similar C23O specific activities were obtained from both wildtype and *lat* mutant backgrounds.

In earlier studies by Yu et al. (234), when the NP1 mutant was complemented by the *lat* gene carried on a high copy number plasmid there was increased cephamycin production and LAT activity but surprisingly the levels of ACVS and IPNS activity also increased. Complementation of the Δ*lat*::*tsr* mutants did not result in increased levels of ACVS or IPNS protein. Complementation of the NP1 mutant with a high copy number plasmid may have resulted in crossing over with the chromosome and this could have elevated the levels of ACVS and IPNS activity, although Piret et al. (181) demonstrated that the complementing plasmid had not stably integrated into the chromosome by Southern blot analysis. However, the fragment used for the complementation studies does not contain the *lat* promoter and therefore the *lat* gene must be transcribed from a vector promoter or via a double crossover. The addition of exogenous αAA increased the level of cephamycin production by the NP1 mutant but unfortunately ACVS and IPNS activities in the αAA supplemented NP1 mutant were not determined. These data could have indicated whether the increased ACVS and IPNS production observed in the complementation experiment resulted from the addition of αAA or from some form of genetic rearrangement. This question could be conclusively answered if the NP1 mutant was transformed with pSET152-based constructs which contain the *lat* gene and would stably integrate into the chromosome at the *attB* site.

In all of the *lat* mutants including the *lat*::*apr* mutant and the NP1 mutant characterized by Yu et al. (234) the ACVS and IPNS activities are reduced to a very similar degree. The classical example of a polar mutation is the *lac* operon in which mutations equally effect the synthesis of all the known components of the system (171). That both of the ACVS and IPNS proteins are reduced to a similar degree suggests that they are part of a single transcript and a mutation which affects the transcript would affect the downstream genes equally.

Surprisingly, the Δ*lat*::*tsr*/term mutants in which transcription of the polycistronic transcript was eliminated by insertion of the FKMT terminator were unable to produce any IPNS protein. Even in the absence of the polycistronic transcript the *pcbC* gene should be transcribed from its own promoter. Petrich et al. (176) suggest that the *pcbC* monocistronic transcript was present in equal amounts to the *lat*/*pcbAB*/*pcbC* polycistronic transcript. Therefore a percentage of the IPNS protein produced should result from translation of the *pcbC* monocistronic transcript. Southern blot analysis of the Δ*lat*::*tsr*/term mutants confirmed that the *pcbC* gene was present and had not been deleted from the chromosome in some secondary mutation. S1 nuclease transcript analysis of RNA isolated from the Δ*lat*::*tsr*/term mutants would be necessary to clarify whether transcription of the *pcbC* monocistronic transcript is occurring in the absence of the polycistronic transcript.

When the *pcbC* gene, cloned into the high copy number promoter probe plasmid pIJ4083, was introduced into *S. lividans* or *S. venezuelae* 13s neither of the resulting transformants were able to make detectable amounts of IPNS activity (175). Replacement of the *pcbC* upstream elements with the *melC1*/*melC* promoter and S/D sequence provided by the plasmid pIJ702 resulted in high level production of active IPNS protein in *S. lividans*. This suggested that the IPNS activity was not produced because the cephamycin non-producing *S. lividans* was unable to recognize the control signals in the *pcbC* promoter necessary to initiate transcription. Promoter probe analysis of the *pcbC*up promoter fragment, unlike the *lat*up promoter fragment, did not suggest a species-specific activation in *S. clavuligerus* and northern blot analysis confirmed that a 1.2 kb *pcbC* transcript was produced by *S. lividans* transformed with the *pcbC* gene promoter probe plasmid construct. Introduction of the *pcbC* gene from *S. griseus* into *S. lividans* also did not result in the production of IPNS activity (84). These results suggested that the *pcbC* gene could not be translated in *S. lividans*.

Expression of the *pcbC* gene in *E. coli* was also a challenge as expression required a two-cistron expression system in which *lacZ'* and the 3' end of *pcbAB* were fused in frame upstream of *pcbC* in the plasmid pUC119 (63). Mutation of the hybrid *lacZ'-pcbAB* cistron caused the fusion protein to prematurely terminate 116 bp upstream of the *pcbC* start codon instead of the normal 31 bp upstream and as a result caused a reduction in IPNS activity by as much as 10-fold. Small regions of dyad symmetry were located in the *pcbAB-pcbC* intergenic region that could form a stem-loop structure that has a strength of -17.2 Kcal/mol, when calculated using the method of Salser (193). This stem-loop structure may occlude the *pcbC* S/D sequence from the ribosome and inhibit translation. The premature termination of the upstream cistron would decrease the translational coupling, as the ribosome translating *lacZ'-pcbAB* would disrupt the secondary structure and promote initiation of *pcbC* translation.

This suggests that the *pcbC* monocistronic transcript would not be translated because of the absence of translation coupling with the upstream *pcbAB* coding region. Therefore the *pcbC* promoter that was active in *S. lividans* created a transcript which could not be translated to produce IPNS protein. This could also explain the lack of IPNS in the Δ*lat*::*tsr*/term mutants as the monocistronic *pcbC* transcript should be present but the mutant could not produce IPNS protein because of the disruption of the *lat*/*pcbAB*/*pcbC* polycistronic transcript. This would also affect the *ccaR* mutants as the elimination of the *lat*/*pcbAB*/*pcbC* polycistronic transcript originating from the CcaR-dependent *lat* promoter would eliminate the production of IPNS protein. This would be independent of whether expression of the *pcbC* promoter is CcaR-dependent since it does not appear that short transcripts originating from the *pcbC* promoter can be translated under the conditions used. The lack of species-specific activity from the *pcbCup* fragment (176) suggests that it would not be CcaR-dependent but the inability of the *pcbC* transcript to be translated explains why no IPNS protein is observed in the *ccaR* mutants.

If this is indeed the case, why does *S. clavuligerus* produce the *pcbC* monocistronic transcript? Originally it was assumed that the monocistronic transcript would allow the cell to make increased IPNS which may be necessary to compensate for degradation of the *pcbC* end of the large multicistronic transcript by 3' ribonucleases. Northern blot analysis of RNA demonstrated a decrease in the amount of degraded high molecular weight mRNA when hybridized with *lat*, *pcbAB* and *pcbC* probes (175). In fact the *pcbC* probe only appears to hybridize with the monocistronic *pcbC* transcript and the only indication of the inclusion of *pcbC* in the polycistronic transcript is the S1 nuclease protection data.

If the monocistronic transcript cannot be translated it suggests that the *pcbC* promoter may be an relic from times when the *lat*, *pcbAB* and *pcbC* were not organized into a tight operon capable of generating a polycistronic message from the strong *lat* promoter. In fact the *pcbC* and *pcbAB* genes are transcribed in opposite orientations in the cephalosporin producing fungi (61, 90) and may not have always had the tight clustering present in *S. clavuligerus* and *N. lactamdurans*.

To demonstrate conclusively that the *pcbC* monocistronic message cannot be translated, a *pcbC* disruption mutant will be transformed with the *pcbC* gene at single copy. The *pcbC* containing fragment will be preceded by a transcription terminator to prevent a translational fusion involving the upstream *pcbAB* coding region from being generated. If the *pcbC* disruption mutant is not complemented by the *pcbC* gene *in trans* and yet the *pcbC* monocistronic transcript is demonstrated to be present in both the "complemented" *pcbC* mutant and in the Δ*lat*::*tsr*/term mutant this would confirm that the *pcbC* monocistronic transcript is not translated. This finding would make its function within *S. clavuligerus* unexplainable and the "interdependence" model impossible.

VI.4 Conclusions

In an effort to understand the genes involved in cephamycin biosynthesis the entire cephamycin cluster except for an internal region within the *pcbAB* gene was sequenced. The *lat*, *pcbAB*, *pcbC*, *cefD*, *cefE* and *cefF* genes along with the newly discovered genes *cmcH*, *cmcI*, *cmcJ* and very likely *pcd* catalyze the biosynthetic steps required to convert lysine, valine and cysteine to cephamycin. The CmcT gene product likely functions as the transporter removing the cephamycin from the cell after production. The PbpA protein likely acts as a resistance mechanism along with the Bla and PcbR proteins to protect the cell from the potentially lethal metabolites that it produces. Roles for the *orf11*, *blp*, *sclU* and *rhsA* in β-lactam biosynthesis are less obvious while the CcaR protein functions as the transcriptional regulator essential for the regulation of cephamycin, clavulanic acid and clavam biosynthesis.

Mutation of the *ccaR* gene eliminated the biosynthesis of cephamycin, clavulanic acid and clavams as well as the production of many of the cephamycin biosynthetic enzymes. Complementation of the *ccaR* mutants returned production of the cephamycin biosynthetic enzymes and the β-lactam antibiotics. CcaR regulated the production of LAT, ACVS and IPNS proteins by controlling transcription of the *lat*/*pcbAB*/*pcbC* polycistronic transcript. The *ccaR* mutants also did not make IPNE and DAOCS which arise from a separate polycistronic transcript, originating from the *cefD* promoter which suggests that the CcaR protein must regulate expression from a number of separate promoters. The elimination of any one of the biosynthetic genes would be sufficient to result in a cephamycin non-producing phenotype but the coordinated control over the biosynthetic genes, all of which may be organized into three separate operons, provides a logical means to regulate antibiotic production and to prevent the wasteful production of biosynthetic enzymes that are not required.

Production of LAT, ACVS and IPNS proteins occurs from translation of the *lat/pcbAB/pcbC* polycistronic transcript originating at the *lat* promoter and ending at the terminator downstream of *pcbC*. Interruption of the polycistronic transcript with a transcription terminator eliminates production of all of the gene products downstream of the terminator. The *pcbC* gene has its own promoter which creates a *pcbC* monocistronic transcript (178) but apparently this transcript can not be translated to produce IPNS protein.

In order to further understand how β-lactam biosynthesis in *S. clavuligerus* is regulated it will be necessary to determine the transcriptional organization of the biosynthetic genes for all of the gene clusters by S1 nuclease protection or northern blot analysis experiments. Comparison of gene expression in wildtype *S. clavuligerus* and *ccaR* mutants will demonstrate which genes are expressed in a CcaR-dependent manner. Comparison of *ccaR* gene expression in *ccaR* mutants and in the NTG 26 and NTG 83 mutants will demonstrate if the CcaR protein regulates its own expression or if a second factor is responsible for regulating its expression.

VII. Bibliography

1. **Adamidis, T., and W. Champness.** 1992. Genetic analysis of *absB*, a *Streptomyces coelicolor* locus involved in global antibiotic regulation. J. Bacteriol. **174**:4622-4628.

2. **Adamidis, T., P. Riggle, and W. Champness.** 1990. Mutations in a new *Streptomyces coelicolor* locus which globally block antibiotic biosynthesis but not sporulation. J. Bacteriol. **172**:2962-2969.

3. **Aharonowitz, Y., G. Cohen, and J. F. Martín.** 1992. Penicillin and cephalosporin biosynthetic genes: structure, organization, regulation and evolution. Annu. Rev. Microbiol. **46**:461-496.

4. **Aharonowitz, Y., and A. L. Demain.** 1978. Carbon catabolite regulation of cephalosporin production in *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. **14**:159-164.

5. **Aidoo, K. A., A. S. Paradkar, D. C. Alexander, and S. E. Jensen.** 1993. Use of recombinant DNA techniques to study the production of β-lactam compounds in *Streptomyces clavuligerus*. p. 219-236. *In* V. P. Gullo, J. C. Hunter-Cevera, R. Cooper, and R. K. Johnson (ed.), Developments in Industrial Microbiology series, vol 33. Society for Industrial Microbiology, Fredricksburg, VA.

6. **Aidoo, K. A., A. Wong, D. C. Alexander, R. A. R. Rittammer, and S. E. Jensen.** 1994. Cloning, sequencing and disruption of a gene from *Streptomyces clavuligerus* involved in clavulanic acid biosynthesis. Gene **147**:41-46.

7. **Alexander, D. C., and S. E. Jensen.** Unpublished results.

8. **Alexander, D. C., L. Lee, and S. E. Jensen.** Unpublished results.

9. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403-410.

10. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389-3402.

11. **Amersham Life Science Inc.** 1993. ECL western blotting protocols. Amersham Life Science Inc., Oakville, ON.

12. **Baggaley, K. H., A. G. Brown, and C. J. Schofield.** 1997. Chemistry and biosynthesis of clavulanic acid and other clavams. Natural Products Reports **14**:309-333.

13. **Bailey, C. R., and D. J. Winstanley.** 1986. Inhibition of restriction in *Streptomyces clavuligerus* by heat treatment. J. Gen. Microbiol. **132**:2945-2947.

14. **Baker, B. J., J. E. Dotzlaf, and W-K. Yeh.** 1991. Deacetoxycephalosporin C hydroxylase of *Streptomyces clavuligerus*. J. Biol. Chem. **266**:5087-5093.

15. **Baldwin, J. E., R. M. Adlington, J. S. Bryans, A. O. Bringhen, J. B. Coates, N. P. Crouch, M. D. Lloyd, C. J. Schofield, S. W. Elson, K. H. Baggaley, R. Cassels, and N. Nicholson.** 1990. Isolation of dihydroclavaminic acid, an intermediate in the biosynthesis of clavulanic acid. Tetrahedron **47**:4089-4100.

16. **Baldwin, J. E., Y. Fujishma, K-C. Goh, and C. J. Schofield.** 1994. Enzymes of valclavam biosynthesis. Tetrahedron Lett. **35**:2783-2786.

17. **Baldwin, J. E., K. D. Merritt, C. J. Schofield, S. W. Elson, and K. H. Baggaley.** 1993. Studies of the stereospecificity of the clavaminic acid synthase catalyzed hydroxylation reaction. J. Chem. Soc. Chem. Commun. 1301-1302.

18. **Barredo, J. L., P. van Soligen, B. Diez, E. Alvarez, J. M. Cantoral, A. Kattevilder, E. B. Smaal, M. A. M. Groenen, A. E. Veenstra, and J. F. Martín.** 1989. Cloning and characterization of acylcoenzymeA:6-aminopenicillanic-acyltransferase gene of *Penicillium chrysogenum*. Gene **83**:291-300.

19. **Barton, B.** Personal communication.

20. **Barton, B., N. Walters, J. Griffin, A. Griffin, and A. Earl.** 1995. Regulation of clavulanic acid and cephamycin biosynthesis in *S. clavuligerus*, abstr. 254, p. 51. *In* Abstracts of the 1995 International Chemical Congress of Pacific Basin Societies (PACIFICHEM '95). Honolulu, HA.

21. **Baylis, H. A., and T. Sinclair.** Personal communication.

22. **Belanger, A. E., G. S. Besra, M. E. Ford, K. Mikusova, J. T. Belisle, P. J. Brennan, and J. M. Inamine.** 1996. The *embAB* genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. Proc. Natl. Acad. Sci. USA. **93**:11919-11924.

23. **Bibb, M. J.** 1996. The regulation of antibiotic production in *Streptomyces coelicolor* A3(2). Microbiol. **142**:1335-1344.

24. **Bibb, M. J., M. J. Bibb, J. M. Ward, and S. N. Cohen.** 1985. Nucleotide sequences encoding and promoting expression of three antibiotic resistance genes indigenous to *Streptomyces*. Mol. Gen. Genet. **199**:26-36.

25. **Bibb, M. J., P. R. Findlay, and M. W. Johnson.** 1984. The relationship between base composition and codon usage in bacterial genes and its use in the simple and reliable identification of protein-coding sequences. Gene **30**:157-166.

26. **Bibb, M. J., G. R. Janssen, and J. M. Ward.** 1985. Cloning and analysis of the promoter region of the erythromycin-resistance gene (*ermE**) of *Streptomyces erythraeus*. Gene **38**:215-226.

27. **Bierman, M., R. Logan, K. O'Brien, E. T. Seno, R. N. Rao, and B. E. Schoner.** 1992. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene **116**:43-49.

28. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **24**:1513-1523.

29. **Blackshear, P. J.** 1984. Systems for polyacrylamide gel electrophoresis. Methods Enzymol. **104**:237-255.

30. **Bradford, M. M.** 1976. A rapid, sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye-binding. Anal. Biochem. **72**:248-254.

31. **Brewer, S. J., P. M. Taylor, and M. K. Turner.** 1980. An adenosine triphosphate-dependent carbamoylphosphate-3-hydroxymethylcephem-*O*-carbamoyltransferase from *Streptomyces clavuligerus*. Biochem. J. **185**:555-564.

32. **Bruton, C. J., E. Guthrie, and K. F. Chater.** 1991. Phage vectors that allow monitoring of transcription of secondary metabolism genes in *Streptomyces*. Bio/Technology **9**:652-656.

33. **Burnham, M. K. R., A. J. Earl, J. H. Bull, D. J. Smith, and G. Turner.** 1989. DNA encoding ACV synthetase. European patent application: 88311655.

34. **Burnham, M. K. R., J. E. Hodgson, and I. D. Normansell.** 1987. Isolation and expression of genes involved in the biosynthesis of beta-lactams. European patent application: 87300757.9.

35. **Busby, R. W., M. D-T. Chang, R. C. Busby, J. Wimp, and C. A. Townsend.** 1995. Expression and purification of two isozymes of clavaminate synthase and initial characterization of the iron binding site. J. Biol. Chem. **270**:4262-4269.

36. **Buttner, M. J., K. F. Chater, and M. J. Bibb.** 1990. Cloning, disruption and transcriptional analysis of three RNA polymerase sigma factor genes in *Streptomyces coelicolor* A3(2). J. Bacteriol. **172**:3367-3378.

37. **Buttner, M. J., and C. J. Lewis.** 1992. Construction and characterization of *Streptomyces coelicolor* A3(2) mutants that are multiply deficient in the nonessential *hrd*-encoded RNA polymerase sigma factors. J. Bacteriol. **174**:5165-5167.

38. **Buttner, M. J., A. J. Smith, and M. J. Bibb.** 1988. At least three different RNA polymerase holoenzymes direct transcription of the agarase gene (*dagA*) of *Streptomyces coelicolor* A3(2). Cell **52**:599-607.

39. **Cattell, K., B. Koop, R. S. Olafson, M. Fellows, I. Bailey, R. W. Olafson, and C. Upton.** 1996. Approaches to detection of distantly related proteins by database searches. BioTechniques **21**:1118-1125.

40. **Chakraburtty, R., and M. J. Bibb.** 1997. The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. J. Bacteriol. **179**:5854-5861.

41. **Chakraburtty, R., J. White, E. Takano and M. J. Bibb.** 1996. Cloning, characterization and disruption of a (p)ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2). Mol. Microbiol. **19**:357-368.

42. **Champness, W. C., and K. F. Chater.** 1994. Regulation and integration of antibiotic production and morphological differentiation in *Streptomyces* spp. p.61-93. *In* P. J. Piggot, C. P. Moran and P. Youngman (ed.), Regulation of Bacterial Differentiation. American Society for Microbiology, Washington, D.C.

43. **Chary, V. K., J. L. de la Fuente, P. Liras, and J. F. Martín.** 1997. *amy* as a reporter gene for promoter activity in *Nocardia lactamdurans*: comparison of promoters of the cephamycin cluster. Appl. Environ. Microbiol. **63**:2977-2982.

44. **Chater, K. F., and M. J. Bibb.** 1997. Regulation of bacterial antibiotic production. p.57-105. *In* H. Kleinkauf and H. von Döhren (ed.), Products of Secondary Metabolism. Biotechnology, vol. 6. VCH, Weinheim, Germany.

45. **Chou, P. Y., and G. D. Fasman.** 1974. Prediction of protein conformation. Biochem. **13**:222-245.

46. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical prediction of protein conformation. Ann. Rev. Biochem. **47**:251-276.

47. **Clayton, T. M., and M. J. Bibb.** 1990. *Streptomyces* promoter-probe plasmids that utilize the *xylE* gene of *Pseudomonas putida*. Nucleic Acids Res. **18**:1077.

48. **Coque, J. J. R., F. J. Enguita, R. E. Cardoza, J. F. Martín, and P. Liras.** 1996. Characterization of the *cefF* gene of *Nocardia lactamdurans* encoding a 3'-methylcephem hydroxylase different from 7-cephem hydroxylase. Appl. Microbiol. Biotechnol. **44**:605-609.

49. **Coque, J. J. R., F. J. Enguita, J. F. Martín, and P. Liras.** 1995. A two-protein component 7α-cephem-methoxylase encoded by two genes of the cephamycin C cluster converts cephalosporin C to 7-methoxycephalosporin C. J. Bacteriol. **177**:2230-2235.

50. **Coque, J. J. R., P. Liras, L. Laiz, and J. F. Martín.** 1993. A gene encoding lysine 6-aminotransferase, which forms the β-lactam precursor α-aminoadipic acid, is located in the cluster of cephamycin biosynthetic genes in *Nocardia lactamdurans*. J. Bacteriol. **173**:6258-6264.

51. **Coque, J. J. R., P. Liras, and J. F. Martín.** 1993. Genes for a β-lactamase, a penicillin-binding protein and a transmembrane protein are clustered with the cephamycin biosynthetic genes in *Nocardia lactamdurans*. EMBO J. **12**:631-639.

52. **Coque, J. J. R., J. F. Martín, J. G. Calzada, and P. Liras.** 1991. The cephamycin biosynthetic genes *pcbAB*, encoding a large multidomain peptide synthetase, and *pcbC* of *Nocardia lactamdurans* are clustered together in an organization different from the same genes in *Acremonium chrysogenum* and *Penicillium chrysogenum*. Mol. Microbiol. **5**:1125-1133.

53. **Coque, J. J. R., J. F. Martín, and P. Liras.** 1993. Characterization and expression in *Streptomyces lividans* of *cefD* and *cefE* genes from *Nocardia lactamdurans*: the organization of the cephamycin gene cluster differs from that in *Streptomyces clavuligerus*. Mol. Gen. Genet. **236**:453-458.

54. **Coque, J. J. R., F. J. Pérez-Llarena, F. J. Enguita, J. L. de la Fuente, J. F. Martín, and P. Liras.** 1995. Characterization of the *cmcH* genes of *Nocardia lactamdurans* and *Streptomyces clavuligerus* encoding a functional 3'-hydroxymethylcephem *O*-carbamoyltransferase from cephamycin biosynthesis. Gene **162**:21-27.

55. **de la Fuente, J. L., A. Rumbero, J. F. Martín, and P. Liras.** 1997. Δ-1-piperideine-6-carboxylate dehydrogenase, a new enzyme that forms α-aminoadipate in *Streptomyces clavuligerus* and other cephamycin C-producing actinomycetes. Biochem. J. **327**:59-64.

56. **Demain, A. L., J. Piret, H. Yu, J. J. R. Coque, P. Liras and J. F. Martín.** 1994. Interdependence of gene expression for early steps of cephalosporin synthesis in *Streptomyces clavuligerus*. Annals N. Y. Acad. Sci. **721**:117-122.

57. **Denhardt, D. T.** 1966. A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Comm. **23**:641-646.

58. **Denis, F., and R. Brzezinski.** 1991. An improved aminoglycoside resistance gene cassette for use in gram-negative bacteria and *Streptomyces*. FEMS Microbiol. Lett. **81**:261-264.

59. **Denis, F., and R. Brzezinski.** 1992. A versatile shuttle cosmid vector for use in *Escherichia coli* and actinomycetes. Gene **111**:115-118.

60. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387-395.

61. **Diez B., S. Gutiérrez, J. L. Barredo, P. van Soligen, L. H. M. van der Voort, and J. F. Martín.** 1990. The cluster of penicillin biosynthetic genes. Identification and characterization of the *pcbAB* gene encoding the (α-aminoadipyl)-L-cysteinyl-D-valine synthetase and linkage to the *pcbC* and *penDE* genes. J. Biol. Chem. **265**:16358-16365.

62. **Doran, J. L., B. K. Leskiw, S. Aippersbach, and S. E. Jensen.** 1990. Isolation and characterization of a β-lactamase inhibitory protein from *Streptomyces clavuligerus* and cloning and analysis of the corresponding gene. J. Bacteriol. **172**:4909-4918.

63. **Doran, J. L., B. K. Leskiw, A. K. Petrich, D. W. S. Westlake, and S. E. Jensen.** 1990. Production of *Streptomyces clavuligerus* isopenicillin N synthase in *Escherichia coli* using two-cistron expression systems. J. Indust. Microbiol. **5**:197-206.

64. **Dotzlaf, J. E., and W-K. Yeh.** 1987. Copurification and characterization of deacetoxycephalosporin C synthetase/hydroxylase from *Cephalosporium acremonium*. J. Bacteriol. **169**:1611-1618.

65. **Durairaj, M., J. L. Doran, and S. E. Jensen.** 1992. High-level expression of the *Streptomyces clavuligerus* isopenicillin N synthase gene in *Escherichia coli*. Appl. Environ. Microbiol. **58**:4038-4041.

66. **Egan, L. A., R. W. Busby, D. Iwata-Reuyl, and C. A. Townsend.** 1997. Probable role of clavaminic acid as the terminal intermediate in the common pathway to clavulanic acid and the antipodal clavam metabolites. J. Am. Chem. Soc. **119**:2348-2355.

67. **Elder, K., and S. E. Jensen.** Unpublished results.

68. **Elson, S. W., K. H. Baggaley, M. Davison, M. Fulston, N. H. Nicholson, G. D. Risbridger, and J. W. Tyler.** 1993. The identification of three new biosynthetic intermediates and one further biosynthetic enzyme in the clavulanic acid pathway. J. Chem. Soc. Chem. Commun. 1212-1214.

69. **Elson, S. W., K. H. Baggaley, J. Gillett, S. Holland, N. H. Nicholson, J. T. Sime, and S. R. Woroniecki.** 1987. Isolation of two novel intracellular β-lactams and a novel dioxygenase cyclising enzyme from *Streptomyces clavuligerus*. J. Chem. Soc. Chem. Commun. 1736-1738.

70. **Enguita, F. J., P. Liras, A. L. Leitão, and J. F. Martín.** 1996. Interaction of the two proteins of the methoxylation system involved in cephamycin C biosynthesis. J. Biol. Chem. **271**:33225-33230.

71. **Feinberg, A., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6-13.

72. **Feinberg, A., and B. Vogelstein.** 1984. Addendum. Anal. Biochem. **137**:266-267.

73. **Feitelson, J. S., F. Malpartida, and D. A. Hopwood.** 1985. Genetic and biochemical characterization of the *red* gene cluster of *Streptomyces coelicolor* A3(2). J. Gen. Microbiol. **131**:2431-2441.

74. **Fernández-Moreno, M. A., J. L. Caballero, D. A. Hopwood, and F. Malpartida.** 1991. The *act* cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* transfer RNA gene of *Streptomyces*. Cell **66**:769-780.

75. **Fernández-Moreno, M. A., A. J. Martín-Triana, E. Martinez, J. Niemi, H. M. Kieser, D. A. Hopwood, and F. Malpartida.** 1992. *abaA*, a new pleiotropic regulatory locus for antibiotic production in *Streptomyces coelicolor*. J. Bacteriol. **174**:2958-2967.

76. **Fleming, A.** 1929. On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenzae*. Br. J. Exp. Pathol. **10**:226.

77. **Floriano, B., and M. J. Bibb.** 1996. *afsR* is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). Mol. Microbiol. **21**:385-396.

78. **Forsman, M., B. Häggström, L. Lindgren, and B. Jaurin.** 1990. Molecular analysis of β-lactamases from four species of *Streptomyces*: comparison of amino acid sequences with those of other β-lactamases. J. Gen. Microbiol. **136**:589-598.

79. **Foster, S. J.** 1993. Molecular analysis of three major wall-associated proteins of *Bacillus subtilis* 168: evidence for processing of the product of a gene encoding a 258 KD precursor two-domain ligand-binding protein. Mol. Microbiol. **8**:299-310.

80. **Foulstone, M., and C. Reading.** 1982. Assay of amoxycillin and clavulanic acid, the components of Augmentin, in biological fluids with high performance liquid chromatography. Antimicrob. Agents Chemother. **22**:753-762.

81. **Fujii, T., E. Takano, H. C. Gramajo, and M. J. Bibb.** 1996. *redD* and *act*II-ORF4, pathway-specific regulatory genes for antibiotic production in *Streptomyces coelicolor* A3(2), are transcribed *in vitro* by an RNA polymerase holoenzyme containing the non-essential σ factor, σ^{hrdD}. J. Bacteriol. **178**:3402-3405.

82. **Fujisawa, Y., and T. Kanazaki.** 1975. Role of acetylCoA: deacetylcephalosporin C acetyltransferase in cephalosporin C biosynthesis by *Cephalosporium acremonium*. Agr. Biol. Chem. **39**:2043-2048.

83. **Furuya, K., and C. R. Hutchinson.** 1996. The DnrN protein of *Streptomyces peucetius*, a pseudo-response regulator is a DNA-binding protein involved in the regulation of daunorubicin biosynthesis. J. Bacteriol. **178**:6310-6318.

84. **García-Domínguez, M., P. Liras, and J. F. Martín.** 1991. Cloning and characterization of the isopenicillin N synthase gene of *Streptomyces griseus* NRRL 3851 and studies of expression and complementation of the cephamycin pathway in *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. **35**:44-52.

85. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods of predicting the secondary structure of globular proteins. J. Mol. Biol. **120**:97-120.

86. **Gramajo, H. C., E. Takano, and M. J. Bibb.** 1993. Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. Mol. Microbiol. 7:837-845.

87. **Guerrero, F. D., J. T. Jones, and J. E. Mullet.** 1990. Turgor-responsive gene transcription and RNA levels increase rapidly when pea shoots are wilted. Sequence and expression of three inducible genes. Plant Mol. Biol. **15**:11-26.

88. **Guilfoile, P. G., and C. R. Hutchinson.** 1992. Sequence and transcriptional analysis of the *Streptomyces glaucescens tcmAR* tetracenomycin C resistance and repressor gene loci. J. Bacteriol. **174**:3651-3658.

89. **Guthrie, E. P., and K. F. Chater.** 1990. The level of a transcript required for production of a *Streptomyces coelicolor* antibiotic is conditionally dependent on a tRNA gene. J. Bacteriol. **172**:6189-6193.

90. **Gutiérrez, S., B. Diez, E. Montenegro, and J. F. Martin.** 1991. Characterization of the *Cephalosporium acremonium pcbAB* gene encoding α-aminoadipylcysteinyl-valine synthetase, a large multidomain peptide synthetase: linkage to the *pcbC* gene as a cluster of early cephalosporin biosynthetic genes and evidence of multiple functional domains. J. Bacteriol. **173**:2354-65.

91. **Gutiérrez, S., J. Velasco, F. J. Fernandez, and J. F. Martín.** 1992. The *cefG* gene of *Cephalosporium acremonium* is linked to the *cefEF* gene and encodes a deacetylcephalosporin C acetyltransferase closely related to homoserine *O*-acetyltransferase. J. Bacteriol. **174**:3056-3064.

92. **Harlow, E., and D. Lane.** 1988. Antibodies: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

93. **Hengge-Aronis, R.** 1993. Survival of hunger and stress: the role of *rpoS* in early stationary phase gene regulation in *E. coli.* Cell:**72**:165-168.

94. **Hill, C. W., C. H. Sandt, and D. A. Vlazny.** 1994. *Rhs* elements of *Escherichia coli*: a family of genetic composites each encoding a large mosaic protein. Mol. Microbiol. **12**:865-871.

95. **Hodgson, J. E., A. P. Fosberry, N. S. Rawlinson, H. N. M. Ross, R. J. Neal, J. C. Arnell, A. J. Earl, and E. J. Lawlor.** 1995. Clavulanic acid biosynthesis in *Streptomyces clavuligerus*: gene cloning and characterization. Gene **166**:49-55.

96. **Hopwood, D. A.** 1981. Genetic studies of antibiotics and other secondary metabolites. p. 187-218. *In* S. W. Glover. and D. A. Hopwood (ed.), Genetics as a Tool in Microbiology (Symposia of the Society for General Microbiology), vol 31. Cambridge University Press, Cambridge.

97. **Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf.** 1985. Genetic manipulation of *Streptomyces*: a Laboratory Manual. John Innes Foundation, Norwich, United Kingdom.

98. **Horinouchi, S., O. Hara, and T. Beppu.** 1983. Cloning of a pleiotropic gene that positively controls biosynthesis of A-factor, actinorhodin and prodigiosin in *Streptomyces coelicolor* A3(2) and *Streptomyces lividans*. J. Bacteriol. **168**:257-269.

99. **Horinouchi, S., M. Kito, M. Nishiyama, K. Furuya, S-K. Hong, K. Miyake, and T. Beppu.** 1990. Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in *Streptomyces coelicolor* A3(2). Gene **95**:49-56.

100. **Ingram, C., M. Brawner, P. Youngman, and J. Westpheling.** 1989. *xylE* functions as an efficient reporter gene in *Streptomyces* spp.: use for the study of *galP1*, a catabolite-controlled promoter. J. Bacteriol. **171**: 6617-6624.

101. **Ishizuka, H., S. Horinouchi, H. M. Kieser, D. A. Hopwood, and T. Beppu.** 1992. A putative two-component regulatory system involved in secondary metabolism in *Streptomyces* spp. J. Bacteriol. **174**:7585-7594.

102. **Jabbouri, S., R. Fellay, F. Talmont, P. Kamalaprija, U. Burger, B. Relic, J. C. Prome, and W. J. Broughton.** 1995. Involvement of *nodS* in N-methylation and *nodU* in 6-*O*-carbamoylation of *Rhizobium* sp. NGR234 nod factors. J. Biol. Chem. **270**:22968-22973.

103. **Janc, J. W., L. A. Egan, and C. A. Townsend.** 1995. Purification and characterization of clavaminate synthase from *Streptomyces antibioticus*. J. Biol. Chem. **270**:5399-5404.

104. **Janssen, G. R., and M. J. Bibb.** 1993. Derivatives of pUC18 that have *Bgl*II sites flanking a modified multiple cloning site and that retain the ability to identify recombinant clones by visual screening of *Escherichia coli* colonies. Gene **124**:133-134.

105. **Jayatilake S, J. A. Huddleston, and E. P. Abraham.** 1981. Conversion of isopenicillin N into penicillin N in cell-free extracts of *Cephalosporium acremonium*. Biochem. J. **194**:645-647.

106. **Jensen, S. E.** Unpublished results.

107. **Jensen, S. E., D. C. Alexander, A. S. Paradkar, and K. A. Aidoo.** 1993. Extending the β-lactam biosynthetic gene cluster in *Streptomyces clavuligerus*, p. 169-176. *In* R. H. Baltz, G. D. Hegeman, and P. L. Skatrud (ed.), Industrial Microorganisms: Basic and Applied Molecular Genetics. American Society for Microbiology. Washington, D.C.

108. **Jensen, S. E., and A. L. Demain.** 1995. β-lactams. p. 239-268. *In* L. C. Vining and C. Stuttard (ed.) Biochemistry and Genetics of Antibiotic Biosynthesis. Butterworth Heinemann, Stoneham, MA.

109. **Jensen S. E., B. K. Leskiw, J. L. Doran, A. K. Petrich, and D. W. S. Westlake.** 1989. Expression of the *Streptomyces clavuligerus* isopenicillin N synthetase gene in *Escherichia coli* and *Streptomyces lividans*. p.239-245. *In* C. L. Hershberger, S. W. Queener, and G. Hegeman (ed.), Genetics and Molecular Microbiology of Industrial Microorganisms. American Society for Microbiology, Washington, D.C.

110. **Jensen, S. E., B. K. Leskiw, L. C. Vining, Y. Aharonowitz, D. W. S. Westlake, and S. Wolfe.** 1986. Purification of isopenicillin N synthetase from *Streptomyces clavuligerus*. Can. J. Microbiol. **32**:953-958.

111. **Jensen, S. E., A. S. Paradkar, and R. H. Mosher.** 1995. Disruption of the gene encoding the lysine ε-aminotransferase in *Streptomyces clavuligerus* and its effect of cephamycin production, abstr. 254, p. 51. *In* Abstracts of the 1995 International Chemical Congress of Pacific Basin Societies (PACIFICHEM '95). Honolulu, HA.

112. **Jensen, S. E., D. W. S. Westlake, R. J. Bowers, and S. Wolfe.** 1982. Cephalosporin formation by cell-free extracts from *Streptomyces clavuligerus*. J. Antibiot. **35**:1351-1360.

113. **Jensen, S. E., D. W. S. Westlake, and S. Wolfe.** 1982. Cyclization of δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine to penicillins by cell-free extracts of *Streptomyces clavuligerus*. J. Antibiot. **35**:483-490.

114. **Jensen, S. E., D. W. S. Westlake, and S. Wolfe.** 1983. Partial purification and characterization of isopenicillin N epimerase activity from *Streptomyces clavuligerus*. Can. J. Microbiol. **29**:1526-1531.

115. **Jensen, S. E., D. W. S. Westlake, and S. Wolfe.** 1985. Deacetoxycephalosporin C synthetase and deacetoxycephalosporin C hydroxylase are two separate enzymes in *Streptomyces clavuligerus*. J. Antibiot. **38**:263-265.

116. **Jensen, S. E., D. W. S. Westlake, and S. Wolfe.** 1988. Production of the penicillin precursor δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV) by cell-free extracts from *Streptomyces clavuligerus*. FEMS. Microbiol. Lett. **49**:213-218.

117. **Jensen, S. E., A. Wong, M. J. Rollins, and D. W. S. Westlake.** 1990. Purification and partial characterization of δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase from *Streptomyces clavuligerus*. J. Bacteriol. **172**:7269-7271.

118. **Jones, G. H., M. S. B. Paget, L. Chamberlin, and M. J. Buttner.** 1997. Sigma-E is required for the production of the antibiotic actnomycin in *Streptomyces antibioticus*. Mol. Microbiol. **23**:169-178.

119. **Kern, B. A., D. Hendlin, and E. Inamine.** 1980. L-lysine ε-aminotransferase involved in cephamycin C synthesis in *Streptomyces lactamdurans*. Antimicrob. Agents Chemother. **17**:676-685.

120. **Kirisawa, R., T. Kobayashi, Y. Kawakami, and H. Iwai.** 1996. Nucleotide sequences of open reading frames 1, 24 and 71 of an attenuated equine herpesvirus-1. J. Equine Sci. **7**:79-87.

121. **Kleinkauf, H., and H. von Döhren.** 1987. Biosynthesis of peptide antibiotics. Ann. Rev. Microbiol. **41**:259-289.

122. **Kovacevic, S., and J. R. Miller.** 1991. Cloning and sequencing of the β-lactam hydroxylase gene (*cefF*) from *Streptomyces clavuligerus*: gene duplication may have led to separate hydroxylase and expandase activities in the actinomycetes. J. Bacteriol. **173**: 398-400.

123. **Kovacevic, S., M. B. Tobin, and J. R. Miller.** 1990. The β-lactam biosynthesis genes for isopenicillin N epimerase and deacetoxycephalosporin C synthetase are expressed from a single transcript in *Streptomyces clavuligerus*. J. Bacteriol. **172**:3952-3958.

124. **Kovacevic, S., B. J. Weigel, M. B. Tobin, T. D. Ingolia, and J. R. Miller.** 1989. Cloning, characterization, and expression in *Escherichia coli* of the *Streptomyces clavuligerus* gene encoding deacetoxycephalosporin C synthetase. J. Bacteriol. **171**:754-760.

125. **Krishnan, H. B., A. Lewin, R. Fellay, W. J. Broughton, and S. G. Pueppke.** 1992 Differential expression of *nodS* accounts for the varied abilities of *Rhizobium fredii* USDA257 and *Rhizobium* sp. strain NGR234 to nodulate *Leucaena* spp. Mol. Microbiol. **6**:3321-3330.

126. **Kuhstoss, S., and R. N. Rao.** 1991. Analysis of the integration function of the streptomycete bacteriophage ØC31. J. Mol. Biol. **222**:897-908.

127. **Kumar, V., J. L. de la Fuente, A. L. Leitão, P. Liras and, J. F. Martín.** 1996. Effect of amplification or targeted disruption of the β-lactamase gene of *Nocardia lactamdurans* on cephamycin biosynthesis. Appl. Microbiol. Biotechnol. **45**:621-628.

128. **Kupka, J., Y-Q. Shen, S. Wolfe, and A. L. Demain.** 1983 Partial purification and properties of the α-ketoglutarate-linked ring-expansion enzyme of β-lactam biosynthesis of *Cephalosporium acremonium*. FEMS Microbiol. Lett. **16**:1-6.

129. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105-132.

130. **Lancini, G., and F. Parenti.** 1982. Antibiotics: an integrated view. Springer-Verlag, New York, NY.

131. **La Roche, S. D., and T. Leisinger.** 1991. Identification of *dcmR*, the regulatory gene governing expression of dichloromethane dehalogenase in *Methylobacterium* sp. strain DM4. J. Bacteriol. **173**:6714-6721.

132. **Lawlor, E. J., H. A. Baylis, and K. F. Chater.** 1987. Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). Genes Dev. **1**:1305-1310.

133. **Lee, P., W. Kuhl, T. Gelbart, T. Kamimura, C. West, and E. Beutler.** 1994. Homology between a human protein and a protein of the green garden pea. Genomics **21**:371-378.

134. **Leskiw, B. K., Y. Aharonowitz, M. Mevarech, S. Wolfe, L. C. Vining, D. W. S. Westlake, and S. E. Jensen.** 1988. Cloning and nucleotide sequence determination of the isopenicillin N synthase gene from *Streptomyces clavuligerus*. Gene **62**:187-196.

135. **Leskiw, B. K., M. J. Bibb, and K. F. Chater.** 1991. The use of a rare codon specifically during development. Mol. Microbiol. **5**: 2861-2867.

136. **Leskiw, B. K., E. J. Lawlor, J. M. Fernandez-Abalos, and K. F. Chater.** 1991. TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. Proc. Natl. Acad. Sci. USA **88**:2461-2465.

137. **Leskiw, B. K., R. Mah, E. J. Lawlor, and K. F. Chater.** 1993. Accumulation of *bldA*-specified tRNA is temporally regulated in *Streptomyces coelicolor* A3(2). J. Bacteriol. **175**:1995-2005.

138. **MacArthur, M. W., and J. M. Thorton.** 1991. Influence of proline residues on protein confromation. J. Mol. Biol. **218**:397-412.

139. **Madduri, K., A. C. DeMarco, R. L. White, C. Stuttard, and L. C. Vining.** 1991. Lysine catabolism and α-aminoadipate synthesis in *Streptomyces clavuligerus*. Appl. Microbiol. Biotechnol. **35**:358-363.

140. **Madduri, K., and C. R. Hutchinson.** 1995. Functional characterization and transcriptional analysis of the $dnrR_1$ locus, which controls daunorubicin biosynthesis in *Streptomyces peucetius*. J. Bacteriol. **177**:1208-1215.

141. **Madduri, K., C. Stuttard, and L. C. Vining.** 1989. Lysine catabolism in *Streptomyces* spp. is primarily through cadaverine: β-lactam producers also make α-aminoadipate. J. Bacteriol. **171**:299-302.

142. **Madduri, K., C. Stuttard, and L. C. Vining.** 1991. Cloning and location of a gene governing lysine ε-aminotransferase, an enzyme initiating β-lactam biosynthesis in *Streptomyces* spp. J. Bacteriol. **173**:985-988.

143. **Malmberg, L-H., W-S. Hu, and D. H. Sherman.** 1993. Precursor flux control through targeted chromosomal insertion of the lysine ε-aminotransferase (*lat*) gene in cephamycin C biosynthesis. J. Bacteriol. **175**:6916-6924.

144. **Marck, C.** 1988. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res. **16**:1829-1836.

145. **Marsh, E. N., M. D. T. Chang, and C. A. Townsend.** 1992. Two isozymes of clavaminate synthase central to clavulanic acid formation: cloning and sequencing of both genes from *Streptomyces clavuligerus*. Biochemistry **31**:12648-12657.

146. **Martín, J. F.** 1992. Cluster of genes for the biosynthesis of antibiotics: regulatory genes and overproduction of pharmaceuticals. J. Indust. Microbiol. **9**:73-90.

147. **Matsumoto, A., S-K. Hong, H. Ishizuka, S. Horinouchi, and T. Beppu.** 1994. Phosphorylation of the AfsR protein involved in secondary metabolism in *Streptomyces* species by a eukaryotic-type protein kinase. Gene **146**: 47-56.

148. **Merrick, M. J.** 1976. A morphological and genetic mapping study of bald colony mutatns of *Streptomyces coelicolor*. J. Gen. Microbiol. **96**:299-315.

149. **Messing, J., and J. Vieira.** 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene **19**:269-276.

150. **Molnár, I., and Y. Murooka.** 1993. Helix-turn-helix DNA-binding motifs of *Streptomyces*: a cautionary note. Mol. Microbiol. **8**:783-787.

151. **Mosher, R. H., and S. E. Jensen.** Unpublished results.

152. **Mosher, R. H., A. S. Paradkar, and S. E. Jensen.** Unpublished results.

153. **Motamedi, H., A. Shafiee, and S-J. Cai.** 1995. Integrative vectors for heterologous gene expression in *Streptomyces* spp. Gene **160**:25-31.

154. **Muller, J. C., V. Toome, D. L. Pruess, J. F. Blount, and M. Weigele.** 1983. Ro 22-5417, a new clavam antibiotic from *Streptomyces clavuligerus*. III. Absolute stereochemistry. J. Antibiot. **36**:217-225.

155. **Narva, K. E., and J. S. Feitelson.** 1990. Nucleotide sequence and transcriptional analysis of the *redD* locus of *Streptomyces coelicolor* A3(2). J. Bacteriol. **172**:326-333.

156. **New England Biolabs Ltd.** 1993. Protein fusion and purification systems. New England Biolabs Ltd., Mississauga, ON.

157. **Newton, G. G. F., and E. P. Abraham.** 1955. Cephalosporin C, a new antibiotic containing sulphur and D-α-aminoadipic acid. Nature **175**:548.

158. **Nicholson, N. H., K. H. Baggaley, R. Cassels, M. Davidson, S. W. Elson, M. Fulston, J. W. Tyler, and S. R. Woroniecki.** 1994. Evidence that the intermediate biosynthetic precursor of clavaulanic acid is its N-aldehyde analogue. J. Chem. Soc. Chem. Commun. 1281-1282.

159. **Niegemann, E., A. Schulz, and K. Bartsch.** 1993. Molecular organization of the *Escherichia coli* gab cluster: nucleotide sequence of the structural genes *gabD* and *gabP* and expression of the GABA permease gene. Arch. Microbiol. **160**:454-460.

160. **Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne.** 1997 Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. **10**:1-6.

161. **Nodwell, J. R., and R. Losick.** 1998. Purification of an extracellular signaling molecule involved in production of aerial mycelium by *Streptomyces coelicolor*. J. Bacteriol. **180**:1334-1337.

162. **Ochi, K.** 1990. A relaxed (*rel*) mutants of *Streptomyces coelicolor* A3(2) with a missing ribosomal protein lacks the ability to accumulate ppGpp, A-factor and prodigiosin. J. Gen. Microbiol. **136**:2405-2412.

163. **Otten, S. L., J. Ferguson, and C. R. Hutchinson.** 1995. Regulation of daunorubicin production in *Streptomyces peucetius* by the *dnrR*$_2$ locus. J. Bacteriol. **177**:1216-1224.

164. **Paradkar, A. S.** Personal communications.

165. **Paradkar, A. S., K. A. Aidoo, and S. E. Jensen.** 1998. A pathway-specific transcriptional activator regulates late steps of clavulanic acid biosynthesis in *Streptomyces clavuligerus.* Mol. Microbiol. **27**:831-843.

166. **Paradkar, A. S., K. A. Aidoo, A. Wong, and S. E. Jensen.** 1996. Molecular analysis of a β-lactam resistance gene encoded within the cephamycin gene cluster of *Streptomyces clavuligerus*. J. Bacteriol. **178**:6266-6274.

167. **Paradkar, A. S., D. C. Alexander, K. A. Aidoo, C. Anders and S. E. Jensen.** 1997. Molecular analysis of clavulanic acid biosynthetic gene cluster in *Streptomyces clavuligerus*, abstr. 0-44, p. 426. *In* Abstracts of the 1997 ASM. Miami, FL.

168. **Paradkar, A. S., and S. E. Jensen.** 1995. Functional analysis of the gene encoding the clavaminate synthase 2 isoenzyme involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. J. Bacteriol. **177**:1307-1314.

169. **Paradkar, A. S., and S. E. Jensen.** Unpublished results.

170. **Paradkar, A. S., S. E. Jensen, and R. H. Mosher.** 1996. Comparative genetics and molecular biology of β-lactam biosynthesis. p.241-277. *In* W. R. Strohl (ed.), Biotechnology of Industrial Antibiotics, second edition. Marcel Dekker, Inc. New York, NY..

171. **Pardee, A. B., F. Jacob, and J. Monod.** 1959. The genetic control and cytoplasmic expression of "inducibility" in the synthesis of β-galactosidase by *E. coli*. J. Mol. Biol. **1**: 165-178.

172. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA. **85**:2444-2448.

173. **Pérez-Llarena, F., J. F. Martín, M. Galleni, J. J. R. Coque, J. L. de la Fuente, J.-M. Frère, and P. Liras.** 1997. The *bla* gene of the cephamycin cluster of *Streptomyces clavuligerus* encodes a class A β-lactamase of low enzymatic activity. J. Bacteriol. **179**:6035-6040.

174. **Pérez-Llarena, F., P. Liras, A. Rodríguez-García, and J. F. Martín.** 1997. A regulatory gene (*ccaR*) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: amplification results in overproduction of both β-lactam compounds. J. Bacteriol. **179**:2053-2059.

175. **Petrich, A. K.** 1993. PhD. thesis. University of Alberta. Transcriptional analysis of the isopenicillin N synthase gene of *Streptomyces clavuligerus.*

176. **Petrich, A. K., B. K. Leskiw, A. S. Paradkar, and S. E. Jensen.** 1994. Transcriptional mapping of the genes encoding the early enzymes of the cephamycin biosynthetic pathway of *Streptomyces clavuligerus*. Gene **142**:41-48.

177. **Petrich, A. K., A. S. Paradkar, and S. E. Jensen.** Unpublished results.

178. **Petrich, A. K., X. Wu, K. L Roy, and S. E. Jensen.** 1992. Transcriptional analysis of the isopenicillin N synthase-encoding gene of *Streptomyces clavuligerus*. Gene **111**:77-84.

179. **Philipp, W. J., S. Poulet, K. Eiglmeier, L. Pascopella, V. Balasubramanian, B. Heym, S. Bergh, B. R. Bloom, W. R. Jacobs, and S. T. Cole.** 1996. An integrated map of the genome of the tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. Proc. Natl. Acad. Sci. USA. **93**:3132-3137.

180. **Pietrusko, R.** 1989. Aldehyde dehydrogenase (EC 1.2.1.3). p.89-127. *In* R. R. Watson (ed.), Biochemistry and Physiology of Substance Abuse, vol 1. CRC Press, Boca Raton, FL

181. **Piret, J., B. Resendiz, B. Mahro, J-Y. Zhang, E. Serpe, J. Romero, N. Connors, and A. L. Demain.** 1990. Characterization and complementation of a cephalosporin-deficient mutant of *Streptomyces clavuligerus* NRRL 3585. Appl. Microbiol. Biotechnol. **32**:560-567.

182. **Pope, M. K., B. D. Green and J. Westpheling.** 1996. The *bld* mutants of *Streptomyces coelicolor* are defective in the regulation of carbon utilization, morphogenesis and cell-cell signalling. Mol. Microbiol. **19**:747-756.

183. **Price N. P. J., B. Relic, F. Talmont, A. Lewin, D. Promé, S. G. Pueppke, F. Maillet, J. Dénarié, J-C. Promé, and W. J. Broughton.** 1992. Broad-host-range *Rhizobium* species strain NGR234 secretes a family of carbamoylated, and fucosylated, nodulation signals that are *O*-acetylated or sulphated. Mol. Microbiol. **6**:3575-3584.

184. **Promega Corp.** 1991. Promega protocols and applications guide. Promega Corp., Madison, WI.

185. **Rao, R. N., M. A. Richardson, and S. Kuhstoss.** 1987. Cosmid shuttle vectors for cloning and analysis of *Streptomyces* DNA. Methods Enzymol. **153**:166-198.

186. **Reading, C., and M. Cole.** 1977. Clavulanic acid: a β-lactamase-inhibiting β-lactam from *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. **11**:852-857.

187. **Rigaud, G., T. Grange, and R. Pictet.** 1987. The use of NaOH as transfer solution of DNA onto nylon membrane decreases the hybridization efficiency. Nucleic Acids Res. **15**:857.

188. **Romero, J., J. F. Martín, P. Liras, A. L. Demain, and N. Rius.** 1997. Partial purification, characterization and nitrogen regulation of the lysine ε-aminotransferase of *Streptomyces clavuligerus*. J. Ind. Microbiol. Biotechnol. **18**:241-246.

189. **Sadosky, A. B., A. Davidson, R. J. Lin, and C. W. Hill.** 1989. *rhs* gene family of *Escherichia coli* K-12. J. Bacteriol. **171**:636-642.

190. Sala-Trepat, J. M., and W. C. Evans. 1971. The meta cleavage of catechol by *Azotobacter* species: 4-oxalocrotonate pathway. Eur. J. Biochem. **20**:400-413.

191. Salowe, S. P., W. J. Krol, D. Iwata-Reuyl, and C. A. Townsend. 1991. Elucidation of the order of oxidations and identification of an intermediate in the multistep clavaminate synthase reaction. Biochem. **30**:2281-2292.

192. Salowe, S. P., E. N. Marsh, and C. A. Townsend. 1990. Purification and characterization of clavaminate synthase from *Streptomyces clavuligerus*: an unusual oxidative enzyme in natural product biosynthesis. Biochem. **29**:6499-6508.

193. Salser, W. 1977. Globin mRNA sequences: analysis of base pairing and evolutionary implications. Cold Spring Harbour Symp. in Quant. Biol. **42**:985-1002.

194. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

195. Samson, S. M., J. E. Dotzlaf, M. L. Slisz, G. W. Becker, R. M. van Frank, L. E. Veal, W. Yeh, J. R. Miller, S. W. Queener, and T. D. Ingolia. 1987. Cloning and expression of the fungal expandase/hydroxylase gene involved in cephalosporin biosynthesis. Bio/Technology **5**:1207-16.

196. Schwecke, T., Y. Aharonowitz, H. Palissa, H. von Döhren, H. Kleinkauf, and H. van Liempt. 1992. Enzymatic characterization of the multifunctional enzyme δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase from *Streptomyces clavuligerus*. Eur. J. Biochem. **205**:687-694.

197. Skatrud, P. L., M. B. Tobin, J. R. Miller, S. Kovacevic, J. Hoskins, S. W. Queener, and T. D. Ingolia. 1994. Evidence for horizontal transfer in the evolution of the β-lactam biosynthetic pathway. p.231-232. In H. Kleinkauf, and H. von Dohren (ed.), 50 Years of Penicillin Application - History and Trends. Public Ltd. Prague, Czechoslovakia.

198. Smith, D. J., M. K. R. Burnham, J. H. Bull, J. E. Hodgson, J. M. Ward. P. Browne, J. Brown, B. Barton, A. J. Earl, and G. Turner. 1990. β-lactam antibiotic biosynthetic genes have been conserved in clusters in prokaryotes and eukaryotes. EMBO **9**:741-747.

199. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503-517.

200. Staskawicz, B., D. Dahlbeck, N. Keen, and C. Napoli. 1987. Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv *glycinea*. J. Bacteriol. **169**:5789-5794.

201. Stroeher, V. L., J. G. Boothe, and A. G. Good. 1995. Molecular cloning and expression of a turgor-responsive gene in *Brassica napus*. Plant Mol. Biol. **27**:541-551.

202. **Strohl, W.** 1992. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. Nucleic Acids Res. **20**:961-974.

203. **Strynadka, N. C. J., S. E. Jensen, P. M. Alzari, and M. N. G. James.** 1996. A potent new mode of β-lactamase inhibition revealed by the 1.7 Å X-ray crystallographic structure of the TEM-1-BLIP complex. Nature Struct. Biol. **3**:290-297.

204. **Strynadka, N. C. J., S. E. Jensen, K. Johns, H. Blanchard, M. Page, A. Matagne, J.-M. Frère, and M. N. G. James.** 1994. Structural and kinetic characterization of a β-lactamase-inhibitor protein. Nature **368**:657-660.

205. **Stuttard, C.** 1982. Temperate phages of *Streptomyces venezuelae*: lysogeny and host specificity shown by phages SV1 and SV2. J. Gen. Microbiol. **128**:115-121.

206. **Stutzman-Engwall, K. J., S. L. Otten, and C. R. Hutchinson.** 1992. Regulation of secondary metabolism in *Streptomyces* spp. and overproduction of daunorubicin in *Streptomyces peucetius*. J. Bacteriol. **174**:144-154.

207. **Takano, E., H. C. Gramajo, E. Strauch, N. Andres, J. White, and M. J. Bibb.** 1992. Transcriptional regulation of the *redD* transcriptional activator gene accounts for growth-phase-dependent production of the antibiotic undecylprodigiosin in *Streptomyces coelicolor* A3(2). Mol. Microbiol. **6**:2797-2804.

208. **Tan, H., and K. F. Chater.** 1993. Two developmentally controlled promoters of *Streptomyces coelicolor* A3(2) that resemble the major class of motility-related promoters in other bacteria. J. Bacteriol. **175**:933-940.

209. **Tang, L., A. Grimm, Y-X. Zhang and C. R. Hutchinson.** 1996. Purification and characterization of the DNA-binding protein DnrI, a transcriptional factor of daunorubicin biosynthesis in *Streptomyces peucetius*. Mol. Microbiol. **22**:801-813.

210. **Tercero, J. A., R. A. Lacalle, and A. Jimenez.** 1993. The *pur8* gene from the pur cluster of *Streptomyces alboniger* encodes a highly hydrophobic polypeptide which confers resistance to puromycin. Eur. J. Biochem. **218**:963-971.

211. **Thirkettle, J. E., J. E. Baldwin, J. Edwards, J. P. Griffin, and C. J. Schoefield.** 1997. The origin of the β-lactam carbons of clavulanic acid. J. Chem. Soc. Chem. Commun. 1025-1026.

212. **Thompson, J. D., D. G. Higgins, and T, J, Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673-4680.

213. **Tiong, S., and D. Nash.** Personal communication.

214. **Tobin, M. B., M. D. Fleming, P. L. Skatrud, and J. R. Miller.** 1990 Molecular characterization of the acyl-coenzyme A:isopenicillin N acyltransferase gene (*penDE*) from *Penicillium chrysogenum* and *Aspergillus nidulans* and activity of recombinant enzyme in *Escherichia coli*. J. Bacteriol. 172:5908-5914.

215. **Tobin, M. B., S. Kovacevic, K. Madduri, J. A. Hoskins, P. L. Skatrud, L. C. Vining, C. Stuttard, and J. R. Miller.** 1991. Localization of the lysine ε-aminotransferase (*lat*) and δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase (*pcbAB*) genes from *Streptomyces clavuligerus* and production of lysine ε-aminotransferase activity in *Escherichia coli*. J. Bacteriol. 173:6223-6229.

216. **Towbin, H., T. Staehelin, and J. Gordon.** 1979 Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350-4354.

217. **Usui, S., and C-A. Yu.** 1989. Purification and properties of isopenicillin N epimerase from *Streptomyces clavuligerus*. Biochim. Biophys. Acta. **999**:78-85.

218. **Valentine, B. P., C. R. Bailey, A. Doherty, J. Morris, S. W. Elson, K. H. Baggaley, and N. H. Nicholson.** 1993. Evidence that arginine is a later metabolic intermediate than ornithine in the biosynthesis of clavulanic acid by *Streptomyces clavuligerus*. J. Chem. Soc. Chem. Commun. 1210-1211.

219. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153**: 3-11.

220. **Waelkens, F., T. Voets, K. Vlassak, J. Vanderleyden, and P. van Rhijn.** 1995. The *nodS* gene of *Rhizobium tropici* strain CIAT899 is necessary for nodulation on *Phaseolus vulgaris* and on *Leucaena leucocephala*. Mol. Plant Microbe Interact. **8**:147-154.

221. **Walters, N. J., B. Barton, and A. J. Earl.** 1994. Novel compounds. International patent: WO 94/18326-A 1.

222. **Ward, J. M., and J. E. Hodgson.** 1993. The biosynthetic genes for clavulanic acid and cephamycin production occur as a 'super-cluster' in three *Streptomyces*. FEMS Microbiol. Lett. **110**:239-242.

223. **Ward, J. M., G. R. Janssen, T. Kieser, M. J. Bibb, M. J. Buttner, and M. J. Bibb.** 1986. Construction and characterisation of a series of multi-copy promoter-probe plasmid vectors from *Streptomyces* using the aminogylcoside phosphotransferase gene from Tn5 as indicator. Mol. Gen. Genet. **203**:468-478.

224. **Wehmeier, U. F.** 1995. New multifunctional *Escherichia coli-Streptomyces* shuttle vectors allowing blue-white screeening on X-Gal plates. Gene **165**:149-150.

225. **White, J., and M. J. Bibb.** 1997. *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. J. Bacteriol. **179**: 627-633.

226. **Whiteman, P. A., E. P. Abraham, J. E. Baldwin, M. D. Fleming, C. J. Schofield, J. D. Sutherland, and A. C. Willis.** 1990. AcylcoenzymeA:6-aminopenicillanic acid acyltransferase from *Penicillium chrysogenum* and *Aspergillus nidulans*. FEBS Lett. **262**:342-344.

227. **Whitney, J. G., D. R. Brannon, J. A. Mabe, and K. J. Wicker.** 1972 Incorporation of labelled precursors into A16886B, a novel β-lactam antibiotic produced by *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. **1**:247-251.

228. **Wietzorrek, A., and M. J. Bibb.** 1997. A novel family of proteins that regulate antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. Mol. Microbiol. **25**:1177-1184.

229. **Wu, T-K., R. W. Busby, T. A. Houston, D. B. McIlwaine, L. A. Egan, and C. A. Townsend.** 1995. Identification, cloning, sequencing, and overexpression of the gene encoding proclavaminate amidino hydrolase and characterization of protein function in clavulanic acid biosynthesis. J. Bacteriol. **177**:3714-3720.

230. **Xiao, X., G. Hintermann, A. Häusler, P. J. Barker, F. Foor, A. L. Demain, and J. Piret.** 1993. Cloning of a *Streptomyces clavuligerus* DNA fragment encoding the cephalosporin 7α-hydroxylase and its expression in *Streptomyces lividans*. Antimicrob. Agents Chemother. **37**:84-88.

231. **Xiao, X., S. Wolfe, and A. L. Demain.** 1991. Purification and characterization of cephalosporin 7-α-hydroxylase from *Streptomyces clavuligerus*. Biochem. J. **280**:471-474.

232. **Ylihonko, K., J. Tuikkanen, S. Jussila, L. Cong, and P. Mäntsälä.** 1996. A gene cluster involved in nogalamycin biosynthesis from *Streptomyces nogalater*: sequence analysis and complementation of early-block mutations in the anthracycline pathway. Mol. Gen. Genet. **251**:113-120.

233. **Yoshida, K., H. Sano, S. Seki, M. Oda, M. Fujimura, and Y. Fujita.** 1995. Cloning and sequencing of a 29 kb region of the *Bacillus subtilis* genome containing the *hut* and *wapA* loci. Microbiol. **141**:337-343.

234. **Yu, H., E. Serpe, J. Romero, J. J. R. Coque, K. Maeda, M. Oelgeschläger, G. Hintermann, P. Liras, J. F. Martín, A. L. Demain, and J. Piret.** 1994. Possible involvement of the lysine ε-aminotransferase gene (*lat*) in the expression of the genes encoding ACV synthetase (*pcbAB*) and isopenicillin N synthase (*pcbC*) in *Streptomyces clavuligerus*. Microbiol. **140**:3367-3377.

235. **Zhang, J., and A. L. Demain.** 1990. Purification of ACV synthetase from *Streptomyces clavuligerus*. Biotechnol. Lett. **12**:649-654.

236. **Zhao, S., C. H. Sandt, G. Feulner, D. A. Vlazny, J. A. Gray, and C. W. Hill.** 1993. Rhs elements of *Escherichia coli* K-12: complex composites of shared and unique components that have different evolutionary histories. J. Bacteriol. **175**:2799-2808.

237. Zukowski, M. M., D. F. Gaffney, D. Speck, M. Kauffmann, A. Findeli, A. Wisecup, and J-P. Lecocq. 1983. Chromogenic identification of genetic regulatory signals in *Bacillus subtilis* based on expression of a cloned *Pseudomonas* gene. Proc. Natl. Acad. Sci. USA. **80**:1101-1105.